May 13, 1999

Perchlorate Peer Review Workshop Report

Prepared for

Office of Solid Waste
U.S. Environmental Protection Agency
401 M St. SW (5307W)
Washington, DC 20460

EPA Contract Number 68-W98-085

RTI Project Number 7200-019

EPA Contract Number 68-W98-085
RTI Project Number 7200-019

Perchlorate Peer Review Workshop Report

May 13, 1999

Prepared for

Office of Solid Waste
U.S. Environmental Protection Agency
401 M St. SW (5307W)
Washington, DC 20460

Prepared by

Center for Environmental Analysis
Research Triangle Institute
Research Triangle Park, NC 27709

Table of Contents

1.0 Summary of Workshop

EPA is in the process of conducting a toxicological review of perchlorate, including the development of a revised provisional Reference Dose (RfD), a cancer assessment, and an ecological assessment. As part of that process, EPA retained the Research Triangle Institute (RTI) to coordinate a scientific peer review of the Draft Toxicological Review Document on perchlorate, entitled "Perchlorate Environmental Contamination: Toxicological Review and Risk Characterization Based on Emerging Information" and a set of supporting toxicological and ecological studies. The peer review was sponsored by EPA's Office of Solid Waste and Emergency Response (OSWER) and Office of Water (OW), while the Draft Toxicological Review Document was prepared by EPA's National Center for Environmental Assessment (NCEA).

In September, 1998, (63 FR 51918) RTI requested nominations from interested stakeholders of highly qualified scientists with expertise in general toxicology, thyroid function and toxicology, developmental toxicology, neurotoxicology, immunotoxicology, pharmacology, genetic toxicology, medical endocrinology with an emphasis on thyroid function, biostatistics, assessment of risks due to non-cancer and cancer health effects, and assessment of risks due to ecological effects.

From a field of outstanding nominees, RTI selected ten independent scientists to be peer reviewers and one of these scientists was selected to chair the workshop. The scientists were selected as peer reviewers based upon their demonstrated expertise in the subject areas outlined above and the need for balance in affiliation among the peer reviewers. At least one expert in each of the above disciplines was selected as a peer reviewer.

The scientists reviewed the following studies:

- A 90-day drinking water toxicity study in rats with ammonium perchlorate
- A neurobehavioral development study of ammonium perchlorate administered orally in drinking water to rats
- Oral (drinking water) dosage-range developmental toxicity study of ammonium perchlorate in rabbits
- Oral (drinking water) developmental toxicity study of ammonium perchlorate in rabbits
- Oral (drinking water) two-generation (one litter per generation) reproduction study of ammonium perchlorate in rats
- Effects of ammonium perchlorate on immunotoxicological, hematological, and thyroid parameters in B6C3F1 mice
- Genotoxicity assays for ammonium perchlorate

- Ecotoxicity assays for ammonium perchlorate

and the draft summary document:

- Perchlorate environmental contamination: Toxicological review and risk characterization based on emerging information.

In January 1999 (64 FR 2492), RTI presented additional information about the workshop. Stakeholders who had additional information on the perchlorate issue were invited to give a short presentation at the workshop.

The peer review workshop was held on February 10 and 11, 1999, in San Bernardino, California. The purpose of the workshop was to provide a forum for the peer reviewers to present and discuss their assessments of the individual studies and the Draft Toxicological Review Document on perchlorate. The public was invited to attend the workshop as observers, and stakeholders were invited to make presentations at the workshop.

Approximately 100 people attended the workshop as observers and 13 stakeholders made short presentations. The first day of the workshop consisted of an overview of the perchlorate issue by EPA, presentations by the stakeholders, and reviews of the individual studies and Draft Toxicological Review Document on perchlorate by the peer reviewers. The second day of the workshop consisted of summaries presented by the peer reviewers on the individual studies and a summary by the Chairman on the issues discussed at the workshop.

This report consists of:

- the Chairman's summary of the conclusions reached by the Peer Review panel at the workshop
- reviews of the studies and Draft Toxicological Review Document on perchlorate prepared by each peer reviewer after the conclusion of the workshop
- the following appendices:
  - Appendix A - the workshop agenda
  - Appendix B - a list of peer reviewers
  - Appendix C - short resumes of peer reviewers
  - Appendix D - a conflict-of-interest statement
  - Appendix E - the charge to the peer reviewers
  - Appendix F - a copy of written comments received by RTI from outside observers before the workshop

- Appendix G - a list of outside observers giving presentations
- Appendix H - a copy of material presented at the workshop by outside observers
- Appendix I - introductory presentations by EPA at the workshop
- Appendix J - a copy of additional analyses from EPA's NCEA.

Some of NCEA's analyses in Appendix J were not available to the peer reviewers prior to the workshop and were not subject to peer review.

2.0 Summary of Findings: Dr. Curtis Klaassen

Peer Review Summary of Perchlorate

HAZARD IDENTIFICATION

This document summarizes the published data, as well as the data that have just been accumulated on perchlorate. We were also informed about further data that are being collected and not yet thoroughly evaluated. In general, the presentation of these results by EPA staff was well done. However, the reviewers want to emphasize that the data that have recently been collected need to be thoroughly evaluated, and that all data should be published in peer-reviewed journals. There was consensus that the thyroid tissue slides from several studies need to be evaluated by a "pathology working group." There was also concern with quality control of TSH measurements across the studies.

There were a number of statistical issues that need to be addressed by the EPA. The major ones are: reexamining all the ANOVAs, making use of all available information in the statistical analyses, considering the elimination of the multiple comparisons correction factor, making a decision on how to consider litter effects, reconsidering the determination of certain no-observed-adverse-effect levels (NOAELs) and lowest-observed-adverse-effect levels (LOAELs), deciding how the benchmark dose calculations are to be used, making certain that the statistical analyses described in the report are the ones actually used, carrying out heterogeneity tests prior to the ANOVAs, and, for reanalyses, comparing the raw data with those of the testing laboratories to make certain they are the same.

Despite these statistical issues it is clear that the numerous studies have delineated the hazards of perchlorate quite nicely. Perchlorate does not appear to produce toxic effects to most organs and is not mutagenic. Insufficient data has been collected to determine its effect on the immune system; however, these studies are in progress.

Toxic effects are produced by perchlorate on the thyroid, affecting T3, T4 and TSH levels and interfering with the developing central nervous system and the thyroid itself. These effects on hormone levels can lead to mental retardation and thyroid tumors. This is an extremely important conclusion, in that it indicates that the toxic effects of perchlorates appear to be limited to the consequence of its inhibition of iodide transport into the thyroid gland.

MODE OF ACTION

The physiology of thyroid hormones has been well studied. Thyroid hormones are important in many basic functions, including regulating basal metabolic rate and development of the central nervous system (as well as other basic functions such as reproduction). If thyroid hormones (T4 and T3) are low, the basal metabolic rate of humans decreases. If the thyroid hormones are low during development of the central nervous system, mental retardation can result, so-called "cretinism" in humans.

Thyroid hormone concentrations are highly regulated. Thus, when thyroid hormones are low, the hypothalamus secretes thyroid releasing hormone (TRH) and the TRH, together with the low levels of thyroid hormones in blood, result in the stimulation of the pituitary to increase its secretion of thyroid stimulating hormone (TSH). The TSH then is carried by the blood to the thyroid gland where it increases many metabolic processes so it can "normalize" the blood levels of thyroid hormones.

The data presented indicate that all of the adverse effects produced by perchlorate are associated with its initial inhibition of iodide uptake to the thyroid gland.

Thyroid Tumors: When perchlorate inhibits the uptake of iodide and production of thyroid hormones, there is a resultant increase in TSH. This adaptive response results in an increase in various enzymes and processes in the thyroid gland, including thyroid hypertrophy. However, if this stimulation of TSH to the thyroid gland is excessive and prolonged, thyroid tumors may result (this effect has been seen in rats). The thyroid tumors are not due to a genotoxic effect, because perchlorate is not mutagenic. These thyroid tumors are due to excessive stimulation of the thyroid gland, resulting in cell proliferation and subsequent neoplasia.

Decreased Mental Development: Thyroid hormones are required for normal mental development. Therefore, when perchlorate inhibits the uptake of iodide by the thyroid gland during the development of the nervous system during the pre- and early post-natal period, it results in a decrease in thyroid hormones in the blood necessary for normal mental development.

It is well accepted that an increase in serum TSH is a more sensitive index of an antithyroid effect and can be observed even if it is not possible to observe a clear decline in serum thyroid hormones (T4 and T3). There is little data to suggest that the increase in serum TSH produces an overcompensation in serum T4 and T3. Therefore the data sets relating to higher serum T3 and T4 levels in ammonium perchlorate-treated rats, such as in the Neurobehavioral Developmental Study, must be interpreted cautiously. In fact, the elevated T4/T3 levels may suggest a problem with the assays or experimental procedures, rather than reflect a real biological phenomenon. Given the current knowledge of the mechanisms of action of perchlorate on the thyroid it seems inappropriate to ascribe an antithyroid or toxic effect to perchlorate in sporadic instances where its treatment is associated with an increase above controls in serum T4 and/or T3.

DOSE RESPONSE

While the hazard identification of perchlorate appears to be in the final stages of being delineated, information on the dose-response is less robust. The panel concluded that thyroid cell hypertrophy (increase in cell size) was not a good biomarker for the adverse effect of perchlorate, but rather suggested the use of hyperplasia (increase in cell number). The present document and the study results often used the terms hyperplasia and hypertrophy interchangeably, which is potentially misleading. Hypertrophy is an adaptive, often reversible response, while hyperplasia is potentially adverse. The EPA also indicated that thyroid tissues from additional animals were available for histopathologic examination. Therefore, it was recommended that a "pathology-working group" be established to look at the existing thyroid slides from the subchronic study and the neurobehavioral study, as well as slides from additional pups in the neurobehavioral developmental toxicity study, to differentiate between hypertrophy and hyperplasia, and to establish dose-response curves for perchlorate-induced hyperplasia.

SPECIAL POPULATIONS

It is well known that thyroid hormones are essential for the normal development of the central nervous system. Therefore, it was suggested that the effects of perchlorate on mental development during fetal development, as well as the first few weeks of life be closely evaluated by histopathological, morphological, and behavioral assessments. Also infant exposure to perchlorate, reduced iodide concentrations, and/or alterations in maternal hormones in the mother's milk should be determined. In addition, it should be ascertained whether perchlorate is transferred in significant dosages from mother to pup via the milk.

RISK ASSESSMENT

***General Issues*:** The classic approach to risk assessment is to determine a NOAEL in laboratory animals, and to divide this by "safety factors" to assure protection of humans. Most of the descriptive studies needed to perform this type of risk assessment on perchlorate have been or are being performed.

A more scientific approach to risk assessment is possible when a greater scientific understanding of the mode and/or mechanism(s) of how the chemical produces the adverse effects has been achieved. With perchlorate, we know the mode of action probably as thoroughly as with any chemical. In addition, pharmacokinetic/toxicokinetic data of perchlorate in rats and humans will be available for building physiologically-based pharmacokinetic (PB-PK) models to decrease the uncertainty of extrapolating data from laboratory animals to humans. Thus, a predictive risk assessment for perchlorate is possible and should be pursued in the next iteration of this assessment.

Risk assessments that estimate human risk from data on human populations should be more accurate than ones that rely on animal data. This can be done with adequate occupational

exposure studies or administration to humans with thyroid disease. With a chemical like perchlorate, which has been used as a drug, a NOEL can be established in humans. The therapeutic dose given to humans is approximately 10 mg/kg. Thus, 10 mg/kg is too high for environmental perchlorate exposure. Further studies on healthy human volunteers are encouraged, such as those presently being conducted at Harvard and in Germany, to establish exposures that will protect humans from potential adverse effects of perchlorate exposure (using as a biomarker a dose that does not increase TSH).

Mode of Action Approach: The document develops a rationale for a mode of action approach for assessing non-cancer and cancer risks of perchlorate. However, the mode of action data were not used in the RfD determinations. A clear statement of EPA's strategy for using mode of action and pharmacokinetic data would be very helpful.

A mode of action approach represents a move in the direction of harmonization of non-cancer and cancer risk assessment and permits a more rational use of available biological and toxicological data. EPA is to be commended for taking a move in the direction of such a harmonized approach.

ECOLOGICAL ASSESSMENT

The ecotoxicology studies were well done and support the screening ecological risk assessment. The major weaknesses of the screening ecological assessment were limited data on exposure and the potential for long-term chronic effects. These limited data resulted in an assessment that was quite conservative in terms of risk-based effect thresholds suggested, and in terms of the scope of additional studies recommended.

CONCLUSIONS

Before the RfD proposed by the EPA can be definitively evaluated, further work is needed. The EPA should convene a pathology working group, review the histopathology from the thyroid slides, and establish dose-response curves for hyperplasia and/or other adverse responses. The variable used by the EPA in the determination of the RfD (thyroid hypertrophy) was not considered by the peer review panel to represent an adverse effect, nor was it demonstrated to be correlated with an adverse response such as hyperplasia. It is not possible at this time to determine what uncertainty factor should applied in the RfD derivation. If the appropriate human data are accumulated, the uncertainty factor could be small. The magnitude of the uncertainty factor needs to wait until all the relevant studies are complete. Based on the lack of demonstrated adverse effects, the RfD proposed by the EPA (0.0009 mg/kg/day) is likely to be conservative.

3.0 Reviews of Toxicity Database, Toxicological Review Document, and Additional Testing Needs by Assigned Panel Members

3.1 General Statistical Issues: Dr. Joseph Haseman

General Comments

While there are a number of statistical issues that require the EPA's attention, all are correctable problems. Dealing with these issues may result in certain changes in the reported NOAELS and/or LOAELS, but it is unlikely that the final conclusions will be substantially altered. Except where noted, these recommendations are in response to Charges 1.1 (3), 1.1 (4), and 1.2.1 (3) from the list of Charges to the External Peer Review Panel. Ten specific and two general recommendations are given below.

Specific Recommendations

Recommendation 1: Re-examine all the ANOVA's, and in those cases where significant interactions and main effects are pooled with error, redo the ANOVA's and Tukey's test pairwise comparisons correctly. When an ANOVA is carried out, the primary objective is to identify important sources of variability and remove these corresponding sums of squares from the error term. These sources of variability may be main effects such as treatment, gender, or replicate, or they may be interactions among these factors.

Once the important sources of variability have been identified and removed from error, the treatment sum of squares is then compared with the error sum of squares to determine if it is significant. What is not appropriate is to identify an important source of variability and then in effect pretend it doesn't exist and add it back in with the error term. This artificially inflates the error and reduces study sensitivity when the treatment sum of squares is compared with this inflated error term. This problem occurs in a number of the EPA statistical analyses.

For example, consider the EPA Analysis of TSH in 90 Day Study at 120 days. From page 41 of Crofton (1998b), the following analysis of variance appears:

	df	SS	MS	F	p
Gender	1	883.7	883.7	310.1	<0.0001
Treatment	3	65.2	21.7	7.63	0.0002
Gender*Treatment	3	12.0	4.0	1.40	0.2497
Error	71	202.3	2.85		

This analysis is fine, and clearly identifies treatment as an important source of variability. However, from page 79 of Crofton (1998b), and Figure 5.9, Page 5-23 of EPA Draft Report, we have

	df	SS	MS	F	p
Treatment	3	65.7	21.9	1.50	0.2224
Error	75	1097.4	14.63		

By inappropriately pooling the gender (and gender*treatment) sum of squares with error, the statistical significance of the treatment effect has been lost. The EPA incorrectly concludes that there is no treatment effect.

For these data there are two possible fixes to this problem: (1) analyze each sex in a separate ANOVA, or (2) analyze the data as eight separate sex/species groups (my choice) . In either case, the Tukey's test comparisons should identify significant treatment effects, especially in females. In fact, these corrected analyses should identify all three doses of females as having shown significant effects. Figure 5.9 should then be re-plotted, showing separate male and female responses, and indicating the significant effects in females.

Similar comments apply to other EPA statistical analyses.

Recommendation 2: Make full use of all available information in the statistical analyses. The two variables that are most critical in the determination of the RfD are thyroid follicular cell hypertrophy and lumen size on PND5. For these variables, the EPA created four (0, 1, 2, 3) or five (0, 1, 2, 3, 4) ordered categories and assigned each animal to one of these categories depending upon the severity of the response.

However, when making pairwise comparisons of each individual dosed group to controls, these ordered responses were essentially ignored, and the data were collapsed into two categories, most often "no response" or "some response."

However, collapsing the data in this way sacrifices information and reduces study sensitivity. After I called this matter to the attention of the EPA, they revised their statistical analysis, and their new analyses seem appropriate and clearly demonstrate the gain in sensitivity by making full use of the study information. Because of the importance of thyroid histology in the RfD determination, these analyses should be brought forward and discussed in the EPA Report itself.

However, further work still needs to be done. For example, the benchmark dose calculations for hypertrophy ignore severity and are based on the collapsed data. The original benchmark dose calculations were also flawed in only making partial use of the data (i.e., using only those animals also having thyroid hormone data). Here again, I recommend making full use of all information available.

During the Peer Review we also learned that there were eight or nine hundred animals in the neurodevelopmental toxicity study that had not yet been examined for thyroid histology. Given the importance of this variable and the small sample sizes (six animals per sex per group) currently used, our panel strongly recommended that the EPA consider including additional animals, especially on PND5, to increase study sensitivity. This matter will be discussed in more detail later, but it is another example of making full use of the information available.

Recommendation 3: Consider eliminating the multiple comparisons correction factor. I understand the principle of why the EPA considered it necessary to use this "correction factor," but at least as currently used, I recommend against it for four reasons. (1) It is not a standard multiple comparisons adjustment. The EPA divides alpha by the square root of the number of variables examined. This is not the widely used Bonferonni adjustment, which divides alpha by the number of variables evaluated. At the very least, the EPA should reference their method of adjustment. (2) This adjustment is used in some studies but not in others. (3) It is reported as being applied to "interaction main effects," a term that is unfamiliar to me (I understand what an interaction is and what a main effect is, but not an interaction main effect. Moreover, this "correction" was reported as being applied to interaction main effects in a study with only one main effect and hence no interactions). (4) It is apparently not applied to the critical pairwise comparisons (Tukey's test), so its impact on study results is minimal in any case.

Recommendation 4: Make a decision about how to consider "litter effects" and be consistent throughout the report. In the contingency table analysis of thyroid follicular cell hypertrophy, the EPA ignores litter and simply reports the response rate among 12 males and females per group. In contrast, in the benchmark dose calculations, these responses are expressed on a per litter basis and an entirely different set of response rates are used for the same data. The EPA should be consistent in their approach to litter effects.

Although I am a strong advocate of the importance of litter effects in other contexts, in this instance I recommend that, for simplicity, the data be reported and evaluated on a per pup basis. There are apparently at most two (but mostly one) pup per litter, so there are insufficient

data to document litter effects if they are present. Moreover, since most litters have only a single pup, there would be little impact of litter effects even if they were present.

If the EPA accepts our panel's recommendation to obtain additional thyroid histology data, then the additional pups could be selected so that each litter had one male and one female. Thus, it would be possible to evaluate, and if necessary adjust for, the impact of litter effects on the response variables of interest.

Incidentally, I do not understand the EPA recommendation in the section on motor activity data to treat gender as a nested factor in an ANOVA. Gender is a crossed factor, not a nested factor. If additional (normally distributed) data are obtained so that there is one male and one female pup per litter, then the appropriate sources of variability for an ANOVA are

Gender
Treatment
Gender x Treatment
Litters (Treatment)
Error.

Litters is a nested factor; gender is not. This section of the draft report requires some work.

Recommendation 5: Reconsider the determination of certain NOAELs and LOAELs. Specific examples include, but may not be limited to:

(a) TSH in the 90 day study: Despite significant elevations in TSH in 0.05 mg/kg-day females on Day 15 (see Figure 5-9, page 5-23), the EPA nevertheless concludes (page 5-22) that "The NOAEL for TSH is 0.05 mg/kg-day." Furthermore, if the Day 120 data are analyzed appropriately, as discussed above, the elevated TSH response in 0.05 mg/kg-day females on Day 120 is also significant by Tukey's test. Thus, I recommend that the EPA lower the NOAEL to at least 0.01 mg/kg-day. Frankly, the data suggest a marginally significant effect at this dose as well, so 0.01 mg/kg-day may in fact be a LOAEL for TSH as well as for T3 and T4.

(b) rT3 in the 14 day study: In this case the EPA reported a NOAEL (0.17 mg/kg-day, see Table 5-2, page 5-14) that was between two doses actually used in the study (0.11-0.12 and 0.44-0.47 mg/kg-day). It is possible that this is just a typographical error, but if it is not, the EPA should explain in more detail how they were able to interpolate a NOAEL between two doses actually used in the study.

(c) Lumen size (morphometric analysis): In this analysis the EPA inappropriately pooled significant block effects with the error term, thereby reducing study sensitivity. I recommend simply comparing each individual dosed group with controls by ANOVA (males and females can be pooled, since there is no gender effect) using the same

basic approach that the EPA employed in the analysis of lumen size in the thyroid histology data. Such an analysis would reveal a significant reduction in lumen size at the 0.1 (but not the 1.0) mg/kg-day dose. Thus, I would conclude that 0.1 mg/kg-day is the LOAEL for lumen size, agreeing with the EPA interpretation for this variable based on the histology data.

However, it is very important to emphasize that even in the revised, more sensitive EPA contingency table analysis, the reduction in lumen size in the (combined male and female) 0.1 mg/kg-day group based on histology is not significant (p=0.0661; see Table 6A of revised Marcus analysis). The EPA was apparently confused by the results of their original statistical analysis and state (page 5-28) that "when data on both sexes were combined, the lowest dose, 0.1 mg/kg-day, was significant at 0.012 (df=8) for the follicular cell hypertrophy and for the lumen size at 0.008 (df=12)." However, these p values are for an overall Chi-square test, and one cannot infer significance at any specific dose based on significant overall differences among groups. These results suggest that the morphometric analyses may be more sensitive than the histology analyses for the detection of lumen size differences, at least at the 0.1 mg/kg-day dose. However, the morphology data had four extra animals per group that were not used in the thyroid histology data analyses.

(d) T4 in Neurodevelopmental Toxicity Study: In my opinion, the T3 and T4 responses track each other very closely in this study (compare Figures 5-11 and 5-12), and although I agree that the reduction in T4 at the 1.0 mg/kg-day dose is not quite statistically significant by Tukey's test, I feel it is biologically significant. Moreover, had a slightly less conservative multiple comparisons test been used (Fisher's protected LSD rather than Tukey's test) the reduction would in fact have been statistically significant. Thus, I see no reason not to consider this dose a LOAEL for T4, and thus 0.1 mg/kg-day a NOAEL for both T3 and T4.

(e) Motor activity data: The EPA seems quite convinced that the motor activity data in high dose males on PND14 show a biologically significant effect, despite (1) the lack of statistical significance at that dose; (2) a trend in the opposite direction in females at PND14; and (3) the lack of any supporting effect in either sex on PND18, PND22, or PND59. In an apparent effort to find statistical significance on PND14, the EPA asked Argus Laboratories to carry out an additional statistical analysis "using gender as a within subject variable, or alternatively, using a nested design with gender nested under litter" (page 5-33). The problems associated with trying to use gender as a nested factor have already been discussed in Recommendation 4 above, and Argus wisely did not attempt such an analysis.

Instead, they responded (York, 1998b) by carrying out an additional statistical analysis using gender as a between-subject variable. Since the females show the opposite effect of males, the overall dose effect is not significant, just as one would

expect (the male effect considered individually is also not significant). The EPA expressed disappointment, noting that the "secondary analysis submitted is still not what the EPA requested" (page 5-34). If the EPA had intended to ask Argus to adjust for litter effects, the analysis they should have requested is summarized in Recommendation 4 above. It is very unlikely that an adjustment for litter effects would have resulted in statistical significance. Indeed, adjusting for litter effects by a "per-litter" analysis generally tends to weaken, not strengthen, treatment effects that might otherwise be detected by a "per-pup" analysis.

The EPA also faults Argus Laboratories for failing to respond adequately to the request for an explanation of why the statistical analysis failed to detect significance in the PND14 motor activity for male rats (page 5-34). The primary reason for the lack of statistical significance is that these data show considerable animal-to- animal variability in motor activity. The EPA Draft report notes this large within group variability (page 5-37), but still concludes that the "increase in activity should be considered biologically significant until additional data can be marshaled to suggest or prove otherwise" (page 5-38). Ultimately, the decision of whether or not a non-significant increase is biologically important is a matter of scientific judgement, but I feel that the alternative explanation given on page 5-37 that this increase "may indeed be a Type I error and would not be found again if this experiment was repeated" is more nearly correct. However, referring to this result as a "Type I error" implies that a statistically significant effect was found, whereas this was not the case for these data.

Another factor that supports my point of view on this matter is that it is the male controls that appear to be unusually low, not the high dose group that is unusually high. For example, the mean motor activity response in the high dose male group (Number of movements = 404; Time spent in movement=364) is very similar to the mean control female response (Number of movements = 393; Time spent in movement = 430; see Argus, Appendix F, pages 509-512). There is nothing in these data that would suggest a gender difference in motor activity. In summary, in my opinion this non-significant increase in motor activity in one dosed group in one sex at one time period simply reflects random variability, nothing more. These comments also respond to Question III A-1 of the charge to the panel.

Recommendation 6: Make a decision about how the benchmark dose calculations are to be used, and make this clear in the report. The report carries out a series of benchmark dose calculations, but it is unclear how the EPA used this information in the RfD determinations. Since the ten page Executive Summary does not even mention the benchmark dose calculations, I can only assume that they played little or no role in the scientific decision making and were possibly carried out "after the fact" to see if such an approach would make a difference. My speculation may or may not be correct, but the EPA needs to address this matter in more detail

and at least mention the benchmark dose calculations in the Executive Summary if they are to remain in the report.

Recommendation 7: Make certain that the statistical analyses described in the report are the ones actually used. In several instances it is not clear what multiple comparison procedure is being used by the EPA. For example, for the morphometric data the actual statistical analyses provided to us (Crofton, 1998f) are Duncan's Multiple Range Test, but Crofton (1998f; page 2 of 29) states that the data were analyzed by "Turkey's Studentized Range Test." This should be "Tukey" rather than "Turkey," but the EPA Draft report also states (page 5-30) that Tukey's test was used. Either of these multiple comparison procedures (Duncan or Tukey) would be okay, but they are different tests. A similar problem occurs in the Segment II Developmental Toxicity studies (and possibly in others; the EPA should review all data analyses).

Similarly, in the original EPA contingency table analyses of the hypertrophy and lumen size data, the EPA Draft Report stated that the Likelihood ratio Chi-square test was being used, whereas in fact the p values reported in some cases were those of the "usual" Chi-square tests.

Recommendation 8: Correct the figures to eliminate the legend for males and females in those cases in which the graph is for one sex only (or for the pool of males and females). A number of figures incorrectly have this legend: Figures 5-2, 5-4, 5-6 (Day 120), 5-9 (Day 90 and Day 120), 5-10, 5-11, 5-12, 5-13, 5-14, 5-15, 5-16, and 5-17. Moreover, for some of these figures, the data should be plotted for males and females separately, since there is a significant gender effect that was ignored by the EPA. The EPA should redo all of these analyses, as indicated earlier, but it appears that at a minimum the following figures should have separate plots for males and females: Figures 5.2, 5-7, and 5.9 (Day 120). Figure 5-8 should also have separate plots for each time period.

Recommendation 9: Carry out heterogeneity tests prior to the ANOVAs and use a variance-stabilizing transformation on the data, if necessary. In the EPA Draft Report and supporting documentation, I could find no mention of the use of heterogeneity tests (e.g., Bartlett's test or Levine's test) prior to the ANOVAs. This is a concern, because an ANOVA assumes homogeneity, and for some variables (e.g., motor activity) it appeared that a data transformation should have been used to equalize the variances. This might also increase study sensitivity.

Recommendation 10: Before performing any reanalyses, compare the raw data with those of the testing laboratory to make certain that they are the same. There were several cases in which the summary statistics from the EPA reanalyses did not agree with the analysis of the same data carried out by the test laboratory.

This led to problems in some cases. For example, in Figure 5-19 for Day 14 (page 5-23), the plotted points for the 0.05 and 0.2 mg/kg-day males have been interchanged (see page 50 of 95 of Crofton, 1998b). This may have occurred in part because the mean responses reported by

the EPA (138.0 and 141.2) differ from the values reported by the test laboratory for these same two groups (140 and 139; see page 447 of SLI Study No. 3455.1).

General Recommendations

Response to Question 1.3.1 (3) charge to the panel. The following two recommendations are even more important than the ten statistical recommendations given above.

Recommendation A: Identify an adverse response upon which to base the RfD. After considerable discussion, it was the final consensus of our panel that thyroid hypertrophy and lumen size should not be considered "adverse responses" by the EPA. The transient or reversible nature of these responses (e.g., gone at the lower doses by PND10; gone from all doses by PND22) also supports this view. Our chairman also expressed the opinion (not challenged by the panel or by the EPA) that neither T3, T4, not TSH should be considered adverse responses from the standpoint of selecting an RfD. Apparently the EPA agrees, since there were statistically significant effects on T3 and T4 at doses as low as 0.01 mg/kg-day that were ignored by the EPA in their RfD calculations.

There apparently is a precedent for the EPA to base an RFD on a response that is not considered adverse, if the non-adverse response can be clearly shown to be predictive of an adverse response at higher doses. However, there is nothing in the EPA report that links the "non-adverse" thyroid hypertrophy response to any adverse response seen at higher doses. Attempts should be made to identify such an adverse response that might be used in the calculation of an RfD. The EPA Draft Report does note that T3, T4, and TSH are inter-correlated and that these variables may also be correlated with thyroid hypertrophy. However, correlating one non-adverse response with other non-adverse responses brings us no closer to identifying a correlated adverse response upon which to base an RfD.

Recommendation B: Carry out a Pathology Working Group (PWG) review of thyroid histopathology, including additional animals on PND5 and possibly other time points if these data are available. One possible adverse response that could be used to determine an RfD is thyroid follicular cell hyperplasia, and a PWG review would ensure that hypertrophy and hyperplasia are clearly defined, diagnosed, and differentiated. In my opinion, a variable this critical should not represent the subjective diagnoses of a single pathologist. I am somewhat surprised that a pathology review is not a standard protocol in EPA laboratory animal studies.

Adding to the possible confusion is the fact that the EPA Draft report and accompanying documentation use the terms hyperplasia and hypertrophy interchangeably (for example, see page 6-2 and also Figure 6-7, page 6-19), whereas they have quite different meanings. If a PWG identifies a dose-related response for thyroid hyperplasia, this variable could be the basis for an RfD rather than thyroid hypertrophy. Our Panel spent little time discussing the various uncertainty factors proposed by the EPA, since such discussions would have little value until an adverse effect can be identified upon which to base an RfD.

Finally, and perhaps most importantly, since no clear adverse effect could be demonstrated in the various studies in the EPA Draft Report (the NOAELs and LOAELs summarized in Tables 6-1 and 6-2 are all for thyroid hormones and thyroid hypertrophy/lumen size, all variables that are considered non-adverse), the RfD reported by the EPA (0.0009 mg/kg-day) is likely conservative.

Minor Comments

(1) In the first panel of Figure 5-9, Day 15 should probably be Day 14 (consistent with the figures for T3 and T4).

(2) There appears to be an error in the raw data for motor activity that requires correction. On page 611 of the Argus 1613-002 submission, the total number of movements for control female 1750 is 108, not 230, as reported in Table F8.

(3) According to page 16 of 19 in Crofton (1998e), the overall F for treatment effects for T3 is 216.89, not 214.89 as indicated in Figure 5-11.

3.2 Neurobehavioral Developmental Study: Dr. R. Thomas Zoeller

Review of Toxicological Review Document

1. In general, the review document appears to do a good job of capturing the key aspects of the protocols and conduct of the studies reviewed. The discussion appears to be balanced, providing a good sense of the strengths and weaknesses of the various studies.

2. In general, the review document appears to adequately evaluate the entire database. Specific issues of weaknesses both in the review document and in the design of studies are discussed below.

3. As written, the review document provides new information on analysis of data presented from the various studies. These new analyses appear reasonable and informative and help improve the relevance of the studies performed.

General Issues

1. Several aspects of the regulation of thyroid hormone "economy" are not fully presented throughout this document, and this weakness is reflected in the construction of experiments in the various studies. These are listed below:

   (a) Page E-4, line 13: "differences in plasma protein binding...." It is true that the half-life of thyroid hormones in circulation is different between rats and humans.

However, it is probably not true that this is simply due to differences in thyroid hormone binding proteins. First, rats do possess the same thyroid hormone binding proteins as expressed in humans (e.g., thyroid binding globulin, TBG, and transthyretin, TTR, [Tani, 1994 #1431]). These circulating proteins are regulated differently; however, pregnant and lactating rats, and neonatal pups, all express TBG in the circulation (Vranckx, 1994 #1432).

(b) Nearly 80% of circulating T_3 is derived from tissue metabolism of T_4. Moreover, tissue metabolism of thyroid hormones is dependent upon iodothyronine deiodinase activities, which are themselves regulated by circulating levels of thyroid hormones (Burmeister, 1997 #1332; Germain, 1997 #1351). Therefore, it is quite important to evaluate the potential effects of perchlorate on tissue metabolism of thyroid hormones, because these effects could be masking effects of perchlorate on the thyroid system.

2. The effect of perchlorate on the thyroid gland is the basis for focusing perchlorate studies on thyroid function and thyroid hormone action. This concept is well-developed in the review document and in the individual studies. However, there is no background information on the role of thyroid hormone in the development of endpoints used for any of the studies. This is a critical issue because thyroid hormone has very specific, and in some cases subtle, effects on the development of the nervous system and endpoints must be designed with this in mind to be valid. At the least, there is no discussion of endpoints that represent valid and sensitive measures of thyroid disruption during development. Moreover, most of the endpoints used do not appear to be sensitive to thyroid disruption and thus, would not be expected to be affected by perchlorate.

3. In discussions of the physiological effects of perchlorate, the document acknowledges that effects are likely to be selectively mediated by the sodium/iodide symporter. Specifically, perchlorate blocks iodide uptake by the symporter. Moreover, the document acknowledges that the symporter is expressed in several regions - not just the thyroid gland - including choriod plexus, gastric and intestinal mucosae, and mammary tissue. What does not appear to be recognized in the review report or in the design of the neurodevelopment studies is the possibility (likelihood), that perchlorate blocks iodide uptake into milk and into the gut and thyroid of the pup. Moreover, it does not appear to recognize the possibility that perchlorate may itself be selectively taken up into milk and transferred to the pup. This is most important in the design of the neurodevelopment studies. However, this issue should be addressed in the review report, especially with respect to the timing and duration of perchlorate exposure.

4. In discussions of perchlorate effects on thyroid hormone synthesis, there appears to be little recognition that the effect of perchlorate on thyroid hormone synthesis will take much longer to develop than the effect on iodide uptake. Because iodinated

thyroglobulin is stored in the colloid, a considerable amount of thyroid hormone precursor is stored. Thus, even if doses of perchlorate are used which completely inhibit iodide uptake, it will still take some time for thyroid hormone levels to respond. This is especially important in the design of the neurodevelopment studies and their interpretation where perchlorate is initiated on the day of conception. In fact, hypothyroxinemia would not be predicted to occur until much later, perhaps as late as parturition.

Specific Comments

1. P3-3, line 10: "This study suggested a LOAEL of 9 mg/kg-day in humans for short-term exposures." This dose is equivalent to 630 mg/day. This seems like a very large dose for a LOAEL when 9.7 mg/kg completely blocks iodide uptake.

2. P3-6, line 1: "There is no information to suggest that humans without Graves' disease would have a similar reaction to perchlorate." It is also true that there is no information to suggest that Graves' patients do not have a similar reaction.

3. Page 4-9, line 26: "...there is a high-affinity binding protein, thyroxine-binding globulin, which binds T_4 (and T_3 to a lesser degree); this protein is missing in rodents and lower vertebrates." This statement is false (Tani, 1994 #1431; Vranckx, 1994 #1432).

4. Is the review document, as currently written, useful for the purpose of characterizing the human health effects of ammonium perchlorate and the perchlorate ion? The database on perchlorate, as described in this review document, does not provide a great deal of confidence in establishing specific human health effects in a developmental context.

5. Development of Reference Dose (RfD)

   (a) The reported NOAELs/LOAELs accurately reflect the findings as reported within the various studies. However, because many of the endpoints were not the most sensitive to thyroid disruption, or analysis methods not the most sensitive, these values are not predicted to be the most accurate in fact. As an important example, the quantitative analysis of thyroid histology performed on pups in the neurodevelopment study were not taken in rigorous manner. Therefore, this LOAEL only reflects the fact that the lowest dose of perchlorate produced an observed adverse effect. If perchlorate treatment were initiated prior to mating, and if the quantitation of thyroid histopathology were performed more rigorously, the LOAEL would undoubtedly be lower than 0.1 mg/kg-day.

(b) The EPA's policy appears to be reasonable and it is appropriate to use the most sensitive endpoint on which to base the RfD. In this case, using the thyroid histopathology study for pups is reasonable because this is a critical and sensitive time for thyroid hormone action and disruption.

(c) The approach is reasonable, as is the identification of the neurodevelopment study. It is important to emphasize that thyroid hormone actions on brain development represent irreversible effects; thus, it is essential to protect the developing brain because there is no therapeutic remedy. However, for reasons discussed above and in the neurodevelopment review, the estimate of 0.1 mg/kg-day is probably high.

(d) Regarding the uncertainty factor of 3 for deriving the RfD: the greatest uncertainty that remains to be determined is the uptake of perchlorate in milk and the potential for concentration of perchlorate in the pup that occurs during lactation. If the assumption about the transfer of perchlorate to the pup is correct, and there is good evidence that it is not, then the uncertainty factor of 3 is reasonable. However, in the absence of these data, an UF of 10 seems justified.

(e) The UF of 3 for using the LOAEL should be increased to 10 because of the issues discussed above for the estimation of this particular LOAEL.

(f) The RfD calculated by the EPA and its justification are quite reasonable. However, the weaknesses in study design and data acquisition as described above introduce caution in this RfD.

Further Testing Needs for Perchlorate

1. It is essential to determine the concentration of perchlorate in milk of animals treated with different doses of perchlorate.

2. Endpoints need to be specifically designed to evaluate thyroid disruption. These are described in detail in the neurodevelopment review.

3. Although it is clear that perchlorate is a thyroid disruptor, the majority of endpoints were not designed to identify thyroid disruption. This is critical, especially considering the potential irreversible and specific effects of thyroid disruption during development. Thus, for an RfD to be developed with confidence, studies on perchlorate-induced disruption of thyroid hormone action are essential and these additional refinements would be important.

3.3 90-Day Subchronic Oral Bioassay Study: Dr. Charles Emerson

This was a study of the potential toxicity of ammonium perchlorate (AP), when ingested in drinking water, in male and female rats. AP was administered to groups of rats in doses ranging from 0 to 10 mg/kg/day for a period 90 days. Most endpoints were evaluated at 14 days and at 90 days as well as after a 30-day recovery period. The parameters evaluated were those relating to thyroid function as reflected by serum thyroid hormones, TSH and thyroid histopathology, indices of reproductive function, mutagenesis as assessed by bone marrow examination for micronuclei formation, routine blood chemistry profiles, and characteristics ascertained by complete gross necropsy and *in vivo* ocular examination. As the workshop progressed, the peer reviewers expressed and adhered to a consensus based on the empirical data and theoretical considerations presented. This consensus was that all of the foreseeable toxicity of perchlorate, and thus its potential toxicity, would be related to its ability to inhibit thyroid function. The potentially deleterious effects of perchlorate, therefore, would be due to the effects of hypothyroidism on tissues and organ systems. In this context the peer reviewers considered the general toxicity parameters evaluated in this study, many of which are not targets for hypothyroidism, to be broad, detailed, and performed in a thorough manner. One reviewer was concerned, however, about the high variability of this data in rats and the potential it held for Type II errors. The assessment of the effects of perchlorate on the thyroid was considered in two contexts. The first was an assessment of thyroid function itself, and the second was the effects of hypothyroidism on organ function and development. It was noted that, despite the major concern being the potential for perchlorate to cause hypothyroidism, the endpoints were not selected on the basis of their being sensitive to hypothyroidism or changes in thyroid status. An important corollary question, one that could not be resolved, is whether rats or humans were more or less sensitive to the same degrees of thyroid hormone deprivation. However, this discussion reinforced that notion that studies in primates might provide the most relevant data concerning the potential toxicity of perchlorate in humans. The most critical aspect of the studies, considering the potential mechanism for perchlorate toxicity, relates to the dose of perchlorate that caused an adverse effect on thyroid function. Studies in humans strongly suggest that a rise in serum TSH is one, if not the most, sensitive and specific index of primary hypothyroidism. A question was raised as to whether the TSH assay used in this study was as sensitive and specific for the rat as TSH assays that use materials (labeled rat TSH, primary antiserum developed against rat TSH, highly purified rat TSH standard) provided by the NIH NIADDK National Hormone and Pituitary Program (Baltimore, MD). It was recommended that serum samples from the 90-Day Subchronic Oral Bioassay Study be tested using the NIH NIADDK rat TSH assay since most research studies in the endocrine literature utilize this assay. As far as hormone assay results themselves were concerned, it was noted that the results for the 10 animals of each sex assumed major importance, as they provided the comparison group for all the treated groups. A larger number of control rats would have been preferable to provide more robust data. Although in some cases, such as was noted for T4 values for males treated for 90 days, relatively low doses of perchlorate were associated with a statistically significant decrease in T4, there was not a clear dose-response relationship that would be expected if the findings were biologically relevant. Only in the case of the 10 mg/kg/day dose at 14 days did the pattern of TSH, T4, and T3 suggest

a meaningful antithyroid effect. Notably, with the exception of the duration of treatment, this is consistent with the conclusion of the AMENDED FINAL REPORT (Joseph C. Siglin, Ph.D. DABT) in which 1.0 mg/kg/day was considered to be the NOAEL for oral administration of ammonium perchlorate to rats via drinking water for 90 days. This NOAEL was based on treatment-related thyroid pathology. In this context the peer reviewers noted that it was important to utilize uniform criteria for thyroid histopathological examination and blinded reading of slides provided the most optimal and scientifically correct way to evaluate samples.

In summary, this is a study of the potential toxicity of ammonium perchlorate when administered to rats for 90 days. The data available at the time of the workshop did not suggest a biologically meaningful adverse effect on a broad range of endpoints at the doses employed except for mild hypothyroidism at a dose of 10 mg/kg/day. It would be important to compare the TSH assay used in this study with that of the NIH NIADDK rat TSH assay. If large discrepancies are observed, strong consideration should be given to assaying all available samples using the NIH NIADDK rat TSH assay. The standards for thyroid histology examination in the 90-Day Subchronic Oral Bioassay Study should be similar to those for other studies and emphasize uniform grading criteria and blinded readings for definitive analysis. In the event of further studies, to confirm or expand the 90-Day Subchronic Oral Bioassay Study, it would be desirable to expand the number of rats in the groups. This would minimize the chances of Type II error and reduce the chance that the mean for the control group is, in fact, an outlier from the other groups.

3.4 Pilot Developmental Study/Segment II Developmental Study: Dr. Rochelle Tyl

I. REVIEW OF INDIVIDUAL STUDIES INITIATED SINCE MAY 1997

A. Oral (Drinking Water) Dosage-Range Developmental Toxicity Study of Ammonium Perchlorate in Rabbits. Final Pilot Report; Protocol No.: 1416-002P. R.G. York, Study Director, Argus Research Laboratories, Inc., Horsham, PA. Report Date: December 10, 1998.

1. Study Design

This is a relatively standard study design for a range-finding study to select target doses for a subsequent "definitive" developmental toxicity study. The design involved five naturally-mated does (presumed pregnant) per group (five groups) to achieve target intakes of ammonium perchlorate of 0, 0.1, 1.0, 10.0, and 20.0 mg/kg/day, in groups 1 through 5, respectively, based on body weights on gestational day (gd) 5 and estimated water consumption of 100 ml/kg/day, with subsequent adjustments based on actual body weights and water consumption values. Correcting the water concentrations, based on the previous week's body weight and actual water consumption values, will always be off

(too low) since the performing laboratory is basing the subsequent week's concentrations on the previous week's body weights, and the animals are gaining weight during the pregnancy. On gd 13, the target concentrations were increased in groups 2 (0.1 mg/kg/day) and 3 (1.0 mg/kg/day) to provide target intakes of 50 and 100 mg/kg/day, respectively, "in order to establish evidence of maternal toxicity." The protocol (Attachment 1) and the text of the report do not provide information on the statistical analyses used. The *ad libitum* exposure period to the does was for gd 6-28, with scheduled sacrifice on gd 29. At sacrifice, the does were examined for body weight, gravid uterine weight, fixed thyroid (and parathyroid) weight, and gross lesions of organs and body cavities. Maternal blood was taken at sacrifice (from the vena cava) for analyses of TSH, T3, and T4 (by AniLytics, Inc., Gaithersburg, MD), and the fixed thyroids were examined histologically by Research Pathology Services, Inc. (New Britain, PA). The number of ovarian corpora lutea and the number and status of uterine implantation sites (total, early, and late resorptions, live and dead fetuses) were recorded. Fetuses were counted, weighed, examined for gross (external) alterations, and sexed internally (one cannot sex rabbit fetuses externally). Fetal internal (visceral) examinations were not performed (except for sex determination), fetal thyroids were not weighed or examined, and fetal blood TSH, T3, or T4 levels were not evaluated. Fetal brains were also not evaluated histologically (a most likely target organ).

2. Conduct of Study

The study was performed under EPA (FIFRA/TSCA) GLPs with a QA statement provided (Attachment 4). The changes in target concentrations and intakes in groups 2 and 3 on gd 13 were done to provide higher intakes to ascertain maternal toxicity (obviously, to that point, the performing laboratory did not see any evidence of maternal toxicity through 20.0 mg/kg/day). This change was documented by Amendment No. 1, Attachment 1, dated December 30, 1997, but the amendment did not specify the calendar date(s) and gd that the change was initiated. (Hopefully, this information is in the study records.) This change makes interpretation of fetal findings at term very problematic, since embryos in groups 2 and 3 were potentially exposed to low doses early in major organogenesis (gd 6-13) and to very high doses (in a "hypothyroid" mother) for the rest of the embryonic (organogenesis) period and throughout the fetal period (see discussion in item no. 5 below). Fetal thyroids should have been examined. The performing laboratory removed one doe at 0 mg/kg/day from study since she had a one-horn pregnancy with only three live fetuses. This is reasonably appropriate (and should be specified in the laboratory's SOPs).

3. Statistical Methodology

Neither the protocol nor the report specifies which statistical tests were used. The protocol (p. 12) only states "Averages and percentages will be calculated. Litter values will be used where appropriate. Additional procedures and/or analyses may be performed

if deemed appropriate." This provides no useful information. I am assuming that nonparametric statistics were employed for continuous (body weights and changes, feed and water consumption, etc.) and discrete (malformation incidences, etc.) data because of the small group size (4-5 does per group), with a $p \leq 0.05$ as standard for statistical significance from the concurrent control group values. The summary data tables also do not specify which statistical tests were used for which parameters.

4. Presentation of Results

This was adequate, except there was no indication of which statistical tests were performed on which parameters. Three fetuses in three different litters (out of five, 60% litter incidence) at 20.0 mg/kg/day exhibited major external malformations. These fetuses were in the group which was exposed to the highest dose during the first half of major organogenesis (the 50 and 100 mg/kg/day groups were not switched to the higher doses until gd 13). Based on the performing laboratory's historical control data (Appendix J to the definitive study final report), meningocele and umbilical hernia (in this study in male fetus 6747-3) were present in low incidences (litters 0.85%, range 0-50%, fetuses 0.11%, range 0-12.0%, for umbilical hernia; and litter 0.24%, range 0-16.7%, fetuses 0.03%, range 0-2.0% for meningocele). Female fetus 6748-3 had cleft palate and a number of other major malformations (only cleft palate in historical control, litter 0.12%, range 0-5.3%; fetuses 0.01%, range 0-0.6%). Male fetus 6749-1 exhibited distended abdomen (ascites), not in historical controls. The performing laboratory acknowledges this could be due to early organogenesis exposure or chance, and correctly suggests that this dose be employed in the definitive study. My copy is missing page 2 and following of Attachment 2, "Thyroid Hormone Levels" from Anilytics, Inc., so I cannot examine individual data for groups 3 through 5.

5. Use in Hazard Characterization

This study was designed, performed, and interpreted to assign target doses (concentrations in the drinking water and intake values) for the subsequent definitive developmental toxicity study. It served this purpose. The changes in doses on gd 13 mean that the does (and the embryos via probable transplacental transport) were exposed to low intakes (0.1 and 1.0 mg/kg/day) for the first half of the period of major organogenesis (gd 6-13), and the does (and embryos/fetuses) were exposed to high intakes (50 and 100 mg/kg/day) for the second half of major organogenesis (gd 13-18/19) and throughout the fetal period (gd 18/19-28). Since the fetal thyroid, at least in humans, begins to concentrate iodide in the second and third trimesters (the fetal period in humans), the embryos/fetuses in these groups potentially were exposed during a critical period of differentiation and prenatal function to high levels of perchlorate. Examination of their thyroids (at least) would have been informative.

6. Incomplete Study

The study is complete, and the final report is available.

B. Oral (Drinking Water) Developmental Toxicity Study of Ammonium Perchlorate in Rabbits. Final Report; Protocol No. 1416-002. R. G. York, Study Director, Argus Research Laboratories, Inc., Horsham, PA. Report Date: September 1, 1998 (final report), September 10, 1998 (report amendment).

1. Study Design

This is a relatively standard study design for a Segment II developmental toxicity study in rabbits. The study design involved 25 naturally-mated does (presumed pregnant) per group (six groups) exposed to the test material *ad libitum* in the drinking water, from gd 6 through 28, to achieve target intakes of ammonium perchlorate of 0, 0.1, 1.0, 10.0, 30.0, and 100.0 mg/kg/day, with adjustments to the concentration of test material based on the previous week's body weights and water consumption. Correcting the water concentrations based on the previous week's body weights and actual water consumption values will always be off (too low), since the performing laboratory is basing the subsequent week's concentrations on the previous week's body weights, and the animals are gaining weight during the pregnancy. Two of the target intakes in this study (10.0 and 30.0 mg/kg/day) bracketed the dose in the range-finding study (20 mg/kg/day) at which fetal malformations were observed. Does were examined and weighed daily, and feed and water consumption was measured daily. Maternal rabbits were sacrificed on gd 29, with blood taken from the inferior vena cava for subsequent analysis for T3, T4, and TSH by AniLytics, Inc. (Gaithersburg, MD). Body weights, gravid uterine weights, and fixed thyroids (plus parathyroids) were weighed, and the thyroids were evaluated histologically by Research Pathology Services, Inc. (New Britain, PA). The number of ovarian corpora lutea and the number and status of uterine implantation sites (total, early and late resorptions, dead and live fetuses) were recorded. Each fetus was weighed, examined grossly, and sexed internally. Live fetuses were euthanized. For approximately 50% of the fetuses per litter, a single cross section was made between the parental and frontal bones of the skull, and the brain was examined *in situ*. The remaining fetuses per litter were decapitated and the heads fixed and decalcified in Bouin's solution for subsequent free-hand serial sections and examination. All fetuses were examined viscerally (abdominal and thoracic organs) and skeletally after staining with alizarin red S (for ossified bone). The report indicates that the skeletal preparations were examined for "skeletal and cartilaginous alterations." A deviation from the protocol and SOPs (Appendix D) indicates that although the protocol and SOPs specify staining with alizarin red S and Alcian blue (the latter specifically for cartilage), "The skeletal specimens were stained with only alizarin red S, not also with Alcian blue." The deviation continues "This deviation did not adversely affect the outcome or interpretation of the study, because although staining with Alcian blue makes it easier to see the cartilage, it is not

necessary to stain the cartilage to evaluate it. Bones and cartilage are evaluated adequately using the single stain, and double staining is not a guideline requirement." Based on many years of experience by this reviewer, it is extremely difficult, if not impossible, to "adequately" assess ossified and cartilaginous skeletal components without double staining (tricks to enhance visualization of cartilage in single-stained specimens include bottom lighting and use of ethanol to make the cartilage "cloudy," but these are not adequate). Fetal thyroids were not weighed or assessed histopathologically, fetal blood was not evaluated for TSH, T3, or T4, and fetal brains (a probable target organ due to maternal and likely fetal hypothyroidism) were not examined histologically.

2. Conduct of Study

This study was performed under EPA (TSCA/FIFRA) GLPs, and there is a QA statement (Appendix K). See comments in answer to question 1 on the major deviation for the protocol on staining fetal skeletal specimens. The lack of staining for fetal cartilaginous skeletal components impairs the laboratory's ability to adequately assess the fetal skeleton at term and identify any possible treatment-related changes.

3. Statistical Methodology

It is adequately described in the protocol (Appendix C, page 14) and in the final report (pages 28-29). However, I have two concerns:

a. The decision tree provided uses the Bartlett's test for homogeneity of variances to decide whether to use parametric or nonparametric statistics. If Bartlett's is significant for a given parameter at $p \leq 0.05$, nonparametric analysis is performed. This sets the "gate" for parametric versus nonparametric analyses very (too) low. Most laboratories, including this reviewer's, use a $p<0.001$ as the gate. I went back and looked at our rabbit data; if we had used the cutoff of $p<0.05$ for Bartlett's, almost all parameters would have been examined nonparametrically. Nonparametric statistics are less robust than parametric statistics, so fewer differences would be identified as statistically significantly different using nonparametric statistics. In my laboratory, we perform an arcsine-square root transformation on all litter-derived percentage data to allow use of parametric methods. For these litter-derived parametric data, we also weigh the ANOVA according to litter size (litters with a larger number of fetuses have lower weight fetuses, relative to litters with fewer numbers of fetuses).

b. Neither the text, summary tables, nor individual tables indicate which tests were used to evaluate statistical significance, so the reviewer has no way of knowing whether parametric or nonparametric tests were used.

4. Presentation of Results

Adequate, although the specific statistical test used is not provided in text or summary tables for a given parameter (see answer to #3).

5. Use in Hazard Characterization

This study is useful and appropriate to use in hazard characterization. The intake values must be corrected for presence of the ammonium ion. Fetal thyroids, brain, and hormone status (TSH, T3, T4) were not evaluated, yet this is most likely the most sensitive life stage, and the brain and thyroid are target organs. The maternal toxicity NOAEL was designated at 1.0 mg/kg/day, due to findings of hypertrophy of the thyroid follicular epithelial cells (and concomitant reduction in follicular lumen size) at 10.0, 30.0, and 100.0 mg/kg/day. The developmental toxicity NOAEL was "greater than 100.0 mg/kg/day," based on the endpoints that were evaluated.

Question: Was a rat developmental toxicity evaluation performed (usually two species, rat and rabbit, are used for developmental toxicity assessments)? At the workshop, we were told there has not been a rat developmental toxicity study. Since the rat is apparently as sensitive as or more sensitive than the human, this is a critical study and an important data gap.

6. Incomplete Study

The study is complete, and the final report is available.

3.5 2-Generation Reproductive Study: Dr. Rochelle Tyl

A. Oral (Drinking Water) Two-Generation (One Litter per Generation) Reproduction Study of Ammonium Perchlorate in Rats. Final Interim Report, Protocol No. 416-001. R. G. York, Study Director. Argus Research Laboratories, Inc., Horsham, PA. Report Date: September 15, 1998.

1. Study Design

This is a two-generation study in CD® (Sprague-Dawley) rats, one litter per generation, performed according to U.S. EPA FIFRA "Pesticide Assessment Guidelines Subdivision F, 83-3" (this is incorrect; it should be 83-4), with additions from the OPPTS draft guidelines (1996), and thyroid (and parathyroid) weights and histopathology for parental and weanling animals. Blood was also taken from parental and weanling animals for "possible analysis of TSH, T3, and T4" (report, p. 17). The study design involved 30 rats/sex/group (four groups) exposed *ad libitum* (summary, p. 16, is incorrect; it states "once daily") to ammonium perchlorate in the

drinking water, to achieve target intakes of 0, 0.3, 3.0, and 30.0 mg/kg/day. Concentrations in the drinking water were adjusted based on the previous week's actual body weight and water consumption values. Correcting the water concentrations, based on the previous week's body weight and actual water consumption values, will always be off (too low), since the performing laboratory is basing the subsequent week's concentrations on the previous week's body weights and the animals are gaining weight rapidly during the F0 and F1 prebreed exposure period and during the F0 and F1 gestational periods. There was also no difference in the concentration provided to the males versus the females, but they differ in their body weights and water consumption and will, therefore, differ (with a fixed concentration the same for both sexes) in their intake in mg/kg/day. There was a ten-week prebreed exposure period (with estrous cyclicity evaluated for the last three weeks), a two-week mating period (based on examination of report Figure 1; not specified in report), three-week gestation for F0 and F1 parents, and three-week lactation period for parents and F1 and F2 offspring. F1 offspring will be evaluated for acquisition of preputial separation (males) beginning on postnatal day (pnd) 39 and of vaginal patency (females) beginning on pnd 28 (too late; CD® rats begin to acquire vaginal patency prior to pnd 28). The current report is an INTERIM report through the necropsy of F1 weanlings; there is no analytical report (Appendix G), no histology report (Appendix H), and no indication that parental and offspring blood will be analyzed for TSH, T3, or T4. At the workshop, additional data were provided on thyroid histopathology of F0 and F1 adult animals, and TSH, T3, and T4 levels for F0 adults, F1 weanlings, and F1 adults at necropsy. It is not clear whether histopathology on F1 weanling thyroids was (or will be) done; it is strongly recommended. It would have been very useful to have brain histology done on F1 and F2 weanlings and to have used the F1 and F2 pups, culled on pnd 4 to standardize litter sizes, for thyroid and brain histopathology (especially to corroborate histopathologic changes in the thyroid observed in pnd 5 pups in the developmental neurotoxicity study), and to evaluate their brains histologically to see if *in utero* exposure in a hypothyroid mother results in neonatal brain alterations. The pituitaries in the F0 and F1 animals also appeared enlarged in a dose-related incidence. This may be indirect evidence that increased TSH synthesis and release occurred in response to recognized reductions in circulating T3 and T4 levels.

2. Conduct of Study

Based on the protocol and data available, it is a well-conducted study, according to current guidelines, with some additional study-specific endpoints added.

3. Statistical Methodology

Statistical analysis is described in the report (pages 37-38) and in the protocol (Appendix D, page 23). I have five concerns with the methodology as presented:

a. It does not provide adequate detail.

b. The use of Bartlett's test for homogeneity of variances to decide whether to use parametric or nonparametric tests with the "gate" at $p<0.05$ (see comments on the rabbit developmental toxicity study) and the implications for robustness of analyses.

c. There is no mention of analysis of covariance (ANCOVA) for day (age) of acquisition of preputial separation (F1 males) or of vaginal patency (females) with body weight at acquisition as the covariate. A number of laboratories running multigeneration studies under the new OPPTS guidelines have data indicating that acquisition of these developmental landmarks is at least partially dependant on body weight, so that lighter pups (due for example to systemic toxicity) will acquire them later.

d. There is no mention of analysis of covariance (ANCOVA) by body weight for anogenital distance at birth of F2 pups if triggered by alterations in reproductive development in F1 animals. Again, anogenital distance is sensitive to body weight.

e. There is no indication of which tests were used to assess statistical significance for a given parameter in the summary tables.

4. Presentation of Results

Based on the interim nature of this report, presentation is adequate. There should be an indication of which statistical test was used for each parameter, parametric or nonparametric.

5. Use in Hazard Characterization

When all the data through weaning of the F2 offspring are in, the thyroid weights and histopathology in parents and offspring, and hopefully parental and offspring TSH, T3, and T4 values, will prove sensitive endpoints to confirm (or refute) the current most sensitive LOAEL of 0.1 mg/kg/day for the pnd 5 pups in the developmental neurotoxicity study. The low dose in this study is 0.3 mg/kg/day, higher than the most sensitive LOAEL, but it will provide confirmation and confidence in the assessment of effects at 0.1 mg/kg/day in the other study. Based on the information provided at the workshop, it is this reviewer's opinion that no NOAEL can be identified, based on thyroid follicular histopathology at 0.3 mg/kg/day (the low dose) in F0 males and females and in F1 males, as well as in both higher doses for both sexes and both generations.

There was much discussion at the Peer Review Workshop on the use of the terms "hypertrophy" (increase in cell size) versus "hyperplasia" (increase in cell number) by the pathologist reading the slides for all three studies; they appeared to be used interchangeably, which is not correct. This has implications for identifying a NOAEL and, therefore, risk assessment, since "hypertrophy" can be reviewed as a transient physiological response within homeostatic constraints (and therefore not necessarily adverse and not a necessary and sufficient biomarker of subsequent hyperplasia and possible neoplasia); i.e., after longer exposure or higher dose). "Hyperplasia" was unanimously viewed by the Panel as adverse and most likely necessary (but not necessarily sufficient) for subsequent neoplasia since cell divisions are involved.

A strong recommendation is to have the slides and diagnoses reviewed by a panel of pathologists to standardize the terminology and confirm the assessments.

6. Incomplete Study

This report is INTERIM and ends at the weaning of the F1 offspring. It does not include analytical results of the drinking water, histopathology of parental and offspring tissues, including thyroid/parathyroid, and it does not include the postnatal growth, development (including reproductive) and mating, and pregnancy and lactational events for F1 parents and F2 offspring. The F1 generation is really the most important, since it was at least potentially exposed during gestation and lactation, with exposures continuing through its mating, gestation and lactation, to produce F2 offspring. Since the fetal hypothalamic-pituitary-thyroid axis may be at especial risk and there are known effects of hypothyroidism on growth, development, and reproduction, this generation may show effects at the low dose of 0.3 mg/kg/day. Since this dose is higher than the low dose for the developmental neurotoxicity study, 0.1 mg/kg/day, at which effects were observed in offspring on pnd 5, it will not alter the most sensitive LOAEL, but it will provide confidence in the assessment.

Additional data made available at the workshop provided thyroid weights and histopathology for F0 and F1 parental animals, and TSH, T3, and T4 levels in F0 and F1 parental animals and F1 weanlings (see comments to no. 5).

B. Review of Toxicological Review Document

1. Effects of Concern to Human Health

Key aspects of study design (protocol), conduct, and conclusions of each toxicology study have been adequately described. Limitations have been appropriately discussed.

2. The strengths of the analyses have been well described.

a. The rat is a good surrogate for humans (rat is at least as sensitive as humans and probably more sensitive).
b. The concept of merging cancer and noncancer endpoints is appropriate, given the early and lower dose histopathologic changes to the thyroid and changes in thyroid hormone homeostasis (TSH, T3, T4) as health risks in and of themselves, and as a predictor ("harbinger") of possible carcinogenic risk. The concern over the use and interpretation of "hypertrophy" versus "hyperplasia" of the thyroid follicular epithelium should at least be addressed and hopefully resolved. Even in the absence of mutagenic potential of perchlorate, the pre- and perinatal animal (and human) should be considered at especial risk.
c. The discussion of the hypothalamic-pituitary-thyroid axis is appropriate.
d. The uncertainty factors are appropriately conservative and protective.

The biggest weakness is the lack of a final report for the two-generation study (only an interim report is available). The F1 generation is crucial and potentially the most sensitive. However, the low dose is 0.3 mg/kg/day, so a lower NOAEL/LOAEL will not be identified.

3. The additional statistical analyses were useful and appropriate. The designation of 0.1 mg/kg/day as a minimal LOAEL, based on the pnd 5 thyroid effects for the developmental neurotoxicology study, depends on whether the findings are "hypertrophy" (increase in cell size, presumably transient and within homeostatic constraints) or "hyperplasia" (increase in cell number, viewed as adverse and a predictor of possible carcinogenic risk since it involves cell divisions). This designation should be in keeping with biological plausibility, mode-of-action, weight of the evidence, and the "precautionary principle."

4. I don't know of any additional references.

5. NA (Ecotox)

6. Under exposure characterization (e.g., page 7-8, line 1 and page 8-6, line 7), the document indicates that "0.37%" is equal to "37x10^6 µg/L." But, 0.37% = 3700 ppm (0.0037 [conversion from percent] x 1,000,000 [1 liter = 1,000,000 µl] = 3700 ppm). Therefore, the correct value must be 3.7x10^6 µg/L (3700 mg/L), at least based on my calculations, which is possible, given the solubility of perchlorate in water of 24.992 w/w % for the ammonium salt (page 2-3, Table 2-2).

7. The document is useful.

C. HAZARD CHARACTERIZATION

Development of Reference Dose (RfD)

1. The individual NOAELS and LOAELS for each study, based on EPA reanalysis, are appropriate.

2. The EPA approach, using a merged cancer and noncancer endpoint evaluation, and viewing changes in histopathology of the thyroid (certainly for hyperplasia, less certain for hypertrophy) and in TSH, T3, and T4 as noncancer endpoints of concern and predictors ("harbingers") of possible cancer risk, is appropriate and justified.

3. The selection of 0.1 mg/kg/day as the "minimum" LOAEL, based on histopathological changes in the thyroid in pups on pnd 5 in the developmental neurotoxicology study, may be appropriate once terminology and diagnoses are standardized and confirmed.

4. Given the nonmutagenicity of perchlorate, the thyroid-pituitary disruptions in homeostasis (histopathology and circulating hormone levels), and the rat as an animal model at least as sensitive as humans (and apparently more sensitive), an uncertain factor of "3" is very conservative to go from rats to humans. A factor of "1" might be more realistic and still protective.

5. The additional uncertainty factor of "3" to go from a LOAEL to a NOAEL is appropriate, given that histopathologic changes are very sensitive and that there were such changes, albeit "minimal" at 0.1 mg/kg/day in the pnd 5 pups (again, depending on the concurrence of "adverse" for hyperplasia and the interpretation of hypertrophy).

6. The use of the RfD for a harmonized human health risk estimate should be protective for both noncancer health effects and cancer endpoints of perchlorate ion. The data support this position.

7. NA (Ecotox)

D. FURTHER TESTING NEEDS FOR PERCHLORATE

Toxicological Testing

1. The experimental designs of the toxicology tests undertaken since May 1997 are adequate, except for:

   a. The rat multigeneration study low dose is 0.3 mg/kg/day, while the low dose and the subsequently identified "minimal" LOAEL in the developmental neurotoxicology study is 0.1 mg/kg/day. "Hindsight is 20-20," but it would

have been very useful to run 0.1 mg/kg/day and a lower dose in the multi-generation study to confirm or refute the 0.1 mg/kg/day LOAEL and identify a true NOAEL in a study examining offspring at least potentially exposed during gestation and lactation, with exposure continuing through their reproductive phase (the F1 to make F2 litters).

b. I hope the blood samples retained from F0 and F1 parental animals and F1 and F2 weanlings for the multi-generation study will be analyzed for TSH, T3, and T4 to accompany the histopathological assessments on the thyroids from these animals. Additional data provided during the workshop included blood samples for F0 and F1 parents and F1 weanlings, and histopathology of F0 and F1 parental organs.

2. There has not been a Segment II developmental toxicity study of ammonium perchlorate in rats (the usual approach, mandated by EPA FIFRA and other regulatory testing requirements, is this test in two species a rodent, usually rat, and a nonrodent, usually a rabbit). It should be done, since the rat is as sensitive as/or more sensitive than the human. *In utero* exposure, up until term, will assess effects of maternal "hypothyroidism" and possible embryo/fetal "hypothyroidism" on *in utero* development, especially in the postembryonic fetal period, which corresponds in humans to when the fetus begins to concentrate iodide into the thyroid (does this indicate activation of the symporter?). The protocol should be a standard OPPTS (1998) developmental toxicity study with maternal blood and maternal (and fetal?) thyroids taken for serum TSH, T3, T4 (dams), and histopathology (dams and fetuses). Fetal brain histopathology should also be done.

3. Development of a PBPK model to address species differences in iodide uptake, perchlorate kinetics, and subsequent perturbations in the hypothalamic-pituitary-thyroid axis is an excellent idea, but:

   a. It will only be as good as the data available and will take considerable time to develop.
   b. It must model maternal, embryo/fetal and perinatal offspring parameters, with more than one timepoint during the pregnancy to detect anticipated changes in the parameters of concern (e.g., iodide uptake, as it changes in the maternal and embryo/fetal compartments during the pregnancy).

4. NA (Ecotox testing)

NOTE: There are a number of typographical errors and more substantive errors in the EPA Toxicological Review (review draft) which will need to be corrected prior to finalization. Examples (not complete) follow:

Review Page	Review Line	Error
E-6	12	Pimephales (fathead minnow) is an aquatic vertebrate (not invertebrate)
1-1	22-23	Thermal explosive decomposition occurs above (not below) 300°C.
3-2	23	"given no [not on] more than..."
3-6	17	"perchlorate" fix typo
3-6	19	"...adverse effect than healthy..."
3-7	17	"...radioimmunoassay..."
4-2	31	"from... to [not and] where..."
4-7	Table 4-1	Superscript "d" and "e" in table but not in footnotes; no superscripts "b" or "c" are present
4-8	8	"...quarternary...", fix typo
4-10	Figure 4-3	Item 9, "TSH-secretory..." [not THS]
4-11	Table 4-2	under "Indirect": (1) "- chemicals inhibiting TH [add "release"], and (2) "- chemical [add "s"] inhibiting [should be stimulating?] hepatic..."
4-18	6	"...developing conceptus..." not "fetus" (too narrow a term)
4-22	Table 4-6	middle column: (a) under rat dev. neurotox, maternal animals are "dams" not does; (b) designate initial parental generation "P0" (dev. neuro.) or "F0" (two-gen), but be consistent
5-43	13	Dev. Tox studies do not employ "brain histology," but serial freehand sections of fixed tissue (no embedment, microtome section or stain, etc.)
5-55	31	"affected" not "effected"
5-56	17	"perchlorate", fix typo

Review Page	Review Line	Error
5-57	30	The description of what the Ames test measures is incorrect. It measures "the reversion from a histidine- (histidine dependent) state to a histidine+ (histidine independent) state induced by chemicals..." (by growing the exposed organism on media without histidine, only his+, independent revertants will grow)
5-59	29	"...drinking water gavage" is incorrect, can't be both; should be "gavage" (I think)
6-10	5	"...diurnal...", fix typo
6-12	9	"...trichloroacetate...", fix typo
6-30	29	"...distributions...", fix typo
6-32	14	"...homeostasis...", fix typo
6-32	21	"...or those treated with anti-thyroid drugs..."; move "anti-thyroid"
6-41	8	"...exquisite...", fix typo
6-45	19	"...because it fell..." (not "if")
6-48	17	"...appropriate...", fix typo
6-52	7	"... because of the efflux..." (delete "to")
7-8, 8-6	1, 7	see comment on "0.37% (37 x 106 µg/L)"
7-14	2-3	If survival is reduced as indicated, it is not dose related; should it read "survival was reduced by 26%...etc."?
9-6	48	"...N-bis (2-hydroxypropyl) nitrosamine...", fix typo

3.6 Immunotoxicity Studies: Dr. Kimber White

Review of Immunotoxicological Studies

Information provided at the Perchlorate Workshop revealed that the laboratory conducting the immunotoxicological evaluation of perchlorate was just starting up and had limited experience in conducting many of the immunological assays. Furthermore, the personnel working on the project was limited to the Principle Investigator, a laboratory technician, and a graduate student. Due to the nature and high visibility of the perchlorate issue, it is unfortunate that a laboratory with more experience in conducting immunological studies was not involved in the project. In addition, based on the workscope undertaken, the personnel resources dedicated to the project do not appear to be sufficient. However, the Principle Investigator has accomplished a significant amount of work with very limited resources, and she should be complemented for her efforts.

The EPA has identified the evaluation of the effects of perchlorate on the immune system as a critical area in the characterization of perchlorate. The question being investigated in these studies (does perchlorate adversely affect the immune system?) is certainly germane to the evaluation of the overall health effects of perchlorate exposure. In general, the experimental design for the immunotoxicology studies appeared appropriate. One of the major strengths of the experimental design followed is the fact that the assays were conducted at multiple time points (14, 90, and 120 days) and, more importantly, each of the assays was conducted at least twice. This approach of repeating all functional assays at least twice is not routinely done in most immunotoxicological investigations, and the fact that the assays were repeated in the evaluation of perchlorate is considered to be a major strength of the study. A second significant strength is the inclusion of the 30-day recovery period (120 day results) in the experimental design. Most immunotoxicological evaluations do not include recovery information on the test compounds. The selection of the B6C3F1 mouse as the mouse strain for use in the immunological investigations is also considered a strength since significant data exist on the immunological responses of this test species. Furthermore, the use of sentinel mice during the study period also represents a strength since the immunological assays are very sensitive to viral infections.

The experimental design also had several weaknesses. First, the decision to use only 6 animals per group is considered to be a weakness. As reflected in the data, on several occasions multiple samples from the same group were lost as a result of technical error or for some other reason, resulting in small group sample size. A decrease in the number of samples in the control or treatment groups can have a significant impact on the statistical evaluation of the results. Another weakness in the design is the fact that a positive control was not included in each of the functional immunological assays. By having a positive control present in the assay, one insures that the assay was conducted correctly and that the assay was capable of detecting an effect, if one was present. The use of a positive control is also helpful as a reference point in evaluating the effect the test compound has on an immunological parameter being measured.

Only one sex, female animals, was evaluated for immunological effects. While this does represent a limitation of the studies, it is consistent with the usual approach used by other organizations, such as the National Toxicology Program, in evaluating the effects of test materials on the immune system. The decision to use female mice and not male mice is considered to be a strength in the design.

The major weakness in the experimental design was the selection of the immunological assays and host resistance assays used in the evaluation of perchlorate. It is unclear from the information provided what the rational was behind the selection of the assays. A tremendous amount of work has gone into developing, validating, and determining the sensitivity and predictability of various immunotoxicological assays (Luster et al., 1988, 1992, and 1993). This work appears not to have been considered in the assay selection for the evaluation of perchlorate. For example, the most predictive immunotoxicological assay, the IgM antibody-forming cell response to the T-dependent antigen, sheep erythrocytes (Plaque Assay), was not originally undertaken in the evaluation of perchlorate.

At the insistence of the EPA, the effect of perchlorate on humoral immunity was evaluated by the immunotoxicology laboratory and the preliminary results from these studies were provided at the workshop. Unfortunately, the assay conducted was not the Plaque Assay for which a significant data base exists, but an ELISA assay to sheep erythrocytes for which only minimum historical data exist. Also, a multilaboratory validation of the ELISA assay has not been conducted. Furthermore, due to the limited number of dilutions utilized in conducting the assays and the manner in which the data were evaluated, the results are essentially unusable for a proper evaluation of the effects of perchlorate on the humoral immune response.

Some of the assays conducted in the evaluation of perchlorate are no longer used by other organizations since they have been found not to be very predictive of immunotoxic effects. In addition, the procedures followed in conducting some of the assays differ significantly from those followed by Luster et al. For example, routinely a Delayed Type Hypersensitivity (DTH) assay, such as those conducted in the Luster et al. studies, consists of a holistic *in vivo* evaluation of the response to the test article. In the Luster et al. DTH studies, a mouse or rat is both sensitized and challenged *in vivo* and the endpoint monitored occurs *in vivo*, such as the swelling of a foot pad or recruitment of cells to the site of challenge. The "DTH" evaluated in these studies is more accurately described as an antigen specific proliferative response as opposed to a "classical" *in vivo* DTH response. As such, it loses the predictive value for identifying potentially immunosuppressive compounds.

It is also unclear what was the rational for selecting and conducting the host resistance assays prior to the completion of the functional assays. If humoral immunity, as evaluated in the plaque assay, is affected to a much greater extent than cell-mediated immune parameters, such as the cytotoxic T-cell response, a host resistance assay designed to evaluate humoral immunity may prove to be more appropriate in the overall immunotoxicological evaluation of perchlorate than the *Listeria monocytogenes* assay which was conducted. Since effects were observed on natural killer cell

activity, the selection of the B16F10 tumor was an appropriate model for evaluation of innate immunity host resistance.

The immunotoxicology studies conducted in the evaluation of perchlorate were reported to have been conducted under the guidelines of Good Laboratory Practices (GLPs). As such, the integrity of the data, from collection until inclusion of tables and graphs in the final report, would have been appropriately monitored by the Quality Assurance Manager and found free of error. Furthermore, the fact that the study was conducted under GLPs indicates that all Standard Operating Procedures (SOPs) for the various assays conducted were in place and were followed in carrying out the study.

A limitation which could decrease the relevance of the study findings is the fact that technical problems occurred with some of the assays. On several occasions the studies' results did not appear to repeat. This may just be due to the fact that the assays were conducted using living biological systems. However, technical problems can also contribute to the failure of an assay to produce reproducible results. The report authors themselves indicated that technical problems occurred in obtaining some of the data (i.e., spleen and thymus cellularity, nitrite production by peritoneal macrophages and *Listeria monocytogenes* study). The combination of the failure of assays to produce reproducible results and the concern raised by the technical problems associated with conducting the assays represents a limitation in the conduct of the study.

All of the immunotoxicological data presented was evaluated by using an analysis of variance and Tukey's multiple comparison to compare control and treatment groups. Results from immunotoxicological assays do not always follow a normal distribution and, at times, a more appropriate evaluation would be one using a non-parametric analysis. Accordingly, this fixed approach used by the authors in evaluating their data is considered to be a weakness. A better procedure would be to use a decision tree approach in which the data are first evaluated to determine if they are normally distributed and then evaluated for homogeneity of variances using the Bartlett's Chi Square Test or another statistical test. Homogeneous data would be evaluated by a parametric one-way analysis of variance. When significant differences occur, treatment groups would be compared to the control using the Dunnett's t Test. Non-homogeneous data would be evaluated using a non-parametric analysis of variance. When significant differences occur, treatment groups would be compared to the control using the Wilcoxon Rank Test.

In general, sufficient data were presented to confirm the findings in the report and usually the data could be tracked from the information provided in the tables to the final results in the figures. Areas of weakness include the lack of necessary information in several of the methods' write-ups. More information needs to be provided on the CD4/CD8 surface marker analysis, such as exactly which antibodies were used and an explanation why other cell types were not included in the evaluation. For example, why were B-cells and total T-cells not counted? Furthermore, the surface marker data need to be presented as absolute numbers as well as in percentage values.

More information needs to be provided on how the Lytic Unit was calculated for the natural killer cell assays. It appears that in some of the natural killer cell assays only three effector-to-target ratios

were used. With such a small number of points defining the curve, accurate determination of Lytic Units is questionable as some of the values have to be extrapolated when 10% lysis of the target cells is used to define the Lytic Unit. Finally, it would be extremely helpful to have all the graphs for a particular assay drawn to the same scale. This would make interpretation across repeat studies and at the different time periods much easier for the reader.

Based on the information provided at the Workshop by the Principle Investigator, including the preliminary ELISA data, as of this time the immunotoxicological evaluation of perchlorate still has significant data "gaps" which prevent us from knowing if perchlorate adversely effects the immune system. While some of the data generated is acceptable and can be used in the evaluation of the effects of perchlorate on components of the immune system, most of the critical information is still not available.

As indicated above, assays which have been shown to be predictive for identifying compounds which are immunotoxic still need to be conducted. As an minimum, the IgM antibody-forming cell response to the T-dependent antigen, sheep erythrocytes (Plaque Assay), needs to be carried out following both 14-and 90-day exposure to perchlorate. Serum from these same animals can also be correctly evaluated in the ELISA to sheep erythrocytes assay. Furthermore, because the *in vitro* phagocytosis assay showed an effect at all dose levels, the effect of perchlorate on macrophage phagocytosis needs to be further evaluated. Preferably, the study should be a holistic *in vivo* evaluation, such as the RES assay which measures the functional ability of the reticuloendothelial system; however, if this is not possible, a more conventional macrophage phagocytosis assay (i.e., using latex particles or radio-labeled sheep erythrocytes) should be conducted. While *Listeria monocytogenes* host resistance studies are currently ongoing, additional host resistance studies may be warranted if the results of the plaque assays identify the humoral immune response as a target for perchlorate exposure. Until these additional data are obtained, the potential for perchlorate exposure to adversely affect the immune system will remain unknown.

3.7 Genotoxicity Studies: Dr. David Brusick

Review and Summary of the Genotoxicity Testing of Perchlorate

Genetic Toxicology

This is a review of individual studies initiated since May 1997 and their impact on the Report.

Strengths and weaknesses of the experimental designs and validity of the interpretations contained in the report: Three studies were included in the series of assays for perchlorate (Salmonella reverse mutation, Mouse Lymphoma assay, and the Mouse Micronucleus). This is an acceptable battery for most chemical screening; however, this array of tests is not as comprehensive as one might want to have for this chemical, considering the extensive human exposure to perchlorate.

Salmonella reverse mutation assay: The assay reported was a very basic design. Most guidelines recommending this test include additional tester strains (e.g., TA104 and/or *E. coli* WP2 *uvr*A). These strains cover mutagenic mechanisms not detected by the initial set of strains and would have strengthened the assay and provided stronger support for assuming that perchlorate is nongenotoxic. Another method which is sometimes used to detect weak agents is the suspension assay which gives better contact between test material and the target organisms. The current assay meets minimum specifications.

Mouse Lymphoma: The design of the study and the conduct of the study were both inadequate. The design should have specific criteria for selecting the high dose on the basis of toxicity. Current guidelines (OECD) require that the high dose be selected at 80-90% toxicity. While it appears that the doses were selected to give the correct toxicity, the actual study data do not provide confirming information of concurrent toxicity. The retest of the -S9 trial should have been set at a 24-hour exposure period based on current protocol requirements. The results from this study are not acceptable for classifying the *in vitro* mutagenicity of perchlorate and should be repeated using current ICH or OECD guidelines.

The Mouse Micronucleus Test: The study design as described in the EPA review and in the contractor report, is incomplete. For example, the design description did not indicate if the mg/kg dose was delivered as a single dose or was given as two or more doses. If the dose was given as a single acute dose (which is what I conclude), there should have been two sample times; one at 24 hours and one at 48 hours. In this study, the sample time (only one) was not stated. I am also concerned about dose selection. The logic used to go from a dose of 2,000 mg/kg, which was moderately toxic, to the high dose in the test of 1,000 mg/kg was not well supported. There was no evidence for toxicity at 1,000 mg/kg, which is contrary to the recommendations from OECD. In order to provide the best possible support for negative genotoxicity, I would have selected a high dose of 2,000 mg/kg (or at least 1,500 mg/kg) and two lower dose levels at 50% and 25% of the high dose. The other potential issue resulting from the dose selection is that one has negative results with no evidence that the test material even reached the target cells (bone marrow). Consequently, the results from this assay are weak support for reaching a conclusion of nongenotoxicity and the study should be repeated with a better design and higher dose levels.

Hazard Characterization

Based on my concerns for at least two of the three tests conducted to define the genetic toxicity of perchlorate, it is my opinion that the classification of perchlorate as nongenotoxic is not well supported. This concern may have some influence on discussions of the type of risk analysis used in future evaluations.

Further Needs for Testing

While the general structure and toxicity of perchlorate do not fit with chemicals associated with DNA activity, I believe that additional testing for DNA damage should be performed including, the repeat of the Mouse Lymphoma assay and the mouse micronucleus assay. Based on completion of those repeats, additional studies may be appropriate.

Review of the Studies Repeated in 1998/1999

Following publication of the draft review report by EPA, all three of the genetic toxicology tests were repeated in different testing facilities.

Salmonella reverse mutation assay: The repeat study conducted by NIEHS contained all of the recommendations cited previously. Strains TA102 and TA104 were added. These strains would respond to active oxygen species or other DNA damaging radicals. In addition, the repeat tests were evaluated using the pre-incubation test methodology. The results of the repeat study were negative confirming the original Ames test.

Mouse Lymphoma: The study was repeated using the current ICH protocol. The toxicity profile seen in the first study was reproduced. The study design included a concurrent toxicity evaluation as part of the main mutation assay. The results of the study were negative and provided support for the negative results reported in the first study.

Mouse Micronucleus Assay: This study was repeated with an intraperitoneal injection route of administration. This route is more sensitive than oral administration (first study) and should have increased the opportunity to see even a weak effect. Dose selection in the repeat study was improved and resulted in treatment doses with evidence of toxicity to the test organism. The results were negative in the repeat study, confirming the results of the first study.

Hazard Characterization

The repeat studies were all well conducted with unequivocal negative results. Following the retests, the genotoxicity section of the EPA review was rewritten incorporating more of the study design information and integrating the two sets of data. The overall interpretation stated in the report was that "mutagenicity is not considered a possible mode of carcinogenic action for this chemical." This interpretation is consistent with the supporting data. The rewrite for this section is much improved over the original.

Summary of Genotoxicity Testing

The genotoxicity of perchlorate was assessed using a battery of three tests:

- The Ames test conducted by plate incorporation and pre-incubation methods.

- The Mouse Lymphoma assay for mutation at the thymidine kinase locus.
- The Mouse Micronucleus assay using both oral and i.p. injection routes of exposure.

All three tests were conducted twice yielding similar negative responses. The negative data from these studies provides strong evidence that perchlorate does not react with DNA to induce either point mutation or chromosome breakage. The EPA review of perchlorate concluded, reasonably, that "mutagenicity is not considered a possible mode of carcinogenic actions for this chemical."

3.8 Ecotoxicity Studies: Dr. Rick Cardwell

Perchlorate Environmental Contamination: Toxicological Review and Risk Characterization Based on Emerging Information

Summary

EPA compiled data on the fate and effects of perchlorate on fish, wildlife, and plants, and from these prepared a screening ecological risk assessment (SERA). These data were sufficient to support preparation of the SERA, and the SERA was prepared in a competent and scientifically defensible manner. The authors demonstrated they were competent with respect to analysis of the ecotoxicological studies and in preparation of ecological risk assessments. The major weaknesses of the SERA reflected limited data concerning (1) exposure and (2) the potential for long-term chronic effects. The limited information resulted in a SERA that was quite conservative in terms of the risk-based effects thresholds suggested and in terms of the scope of additional studies recommended. Defining the concentrations of perchlorate that occur in surface waters, sediments, soils, and the tissues of plants and animals clearly is the most important data gap currently. The current information on exposure is so meager that it is impossible to judge the ecological problem posed by perchlorate, if one exists. The lack of this information makes it impossible to decide what types of fish, wildlife, and plants are at risk. The latter is important for making the risk assessments more specific and to guide testing needs. It is assumed that ponds, wetlands, and irrigation canals that are subject to evapoconcentration and occur downgradient from known perchlorate sources may be at greatest risk. If so, the fish, wildlife, and plants occurring in these habitats may be at greatest risk. Because of the possibly significant bioaccumulation of perchlorate that has been documented in leaves of some plants, effects to wildlife from consumption of perchlorate via dietary pathways may warrant special emphasis. However, bioaccumulation in another food base, invertebrates, should be examined, for both water-based and food-based exposures (e.g., insects consuming vascular plants or detritus). And finally, long-term chronic effects need examination in species and endpoints sensitive to the thyroid hormone. There have been no long-term studies of this type in species other than laboratory rodents (rat). The rat studies supporting human health risk evaluations should suffice for ecological risk evaluations made for mammals, but not for birds, fish, and invertebrates. The critical ecotoxicological studies may include the following. A long-term chronic exposure with fathead minnows encompassing gametogenesis and embryogenesis up

through to 28-day old fry is one such study. A study of the effects on tadpole development up through metamorphosis would be helpful in deciding risks to amphibians, as the frog embryo development study is inconclusive with respect to hormone-mediated effects. And finally, a chronic study using waterfowl should be considered if the foregoing rat, fathead minnow, and amphibian studies suggest effects at concentrations significantly lower than those currently predicted with the ecological screening risk analysis.

Specific Comments

***Ecotoxicological Effects of Concern* (Manuscript by Sprenger et al. (1998)**

Have the Key Aspects of Protocols, Methods and Ecotox Study Results Been Described Adequately?

EA Engineering, Science and Technology, Inc.: EA conducted the following tests: fathead minnow (*Pimephales*) and *Daphnia magna* acutes, fathead minnow 7-day chronic, *Ceriodaphnia* 7-day acute and chronic and 28-day lettuce chronic. The test protocols and methods used by EA Engineering, Science and Technology have been adequately described; more importantly, they conform to standard methods (e.g., ASTM and EPA). The test results meet the protocol requirements and are of good quality. The range-finding and definitive test results are in reasonable agreement, lending further confidence to the validity of the definitive test results.

Block Environmental Services, Inc.: BES conducted 6-7 day chronics with fathead minnow and *Ceriodaphnia* in essentially the same manner as EA. The protocol and test documentation were significantly less extensive than those of EA. No protocol was provided, although BEA said it was available and similar to methods described in a EPA chronic test method manual. Only data summaries were provided, and the results could not be independently verified. However, the *Ceriodaphnia* test results were very close to those reported by EA, and the *Pimephales* data were within a range (factor of 3-4) that is usual for inter-laboratory variability. This study is believed to be of acceptable quality.

Dumont and Bantle (1998): Drs. Dumont and Bantle studied mortality and malformations in the African toad *Xenopus* exposed to aqueous perchlorate. Although a formal test protocol did not accompany the report, it was performed according to ASTM and contained extensive test data. These data appear to be of acceptable quality.

Nzengung (1998): Dr. Nzengung studied the uptake and metabolism of perchlorate in three woody plants (willow – *Salix*; Eastern cottonwood – *Poplar*; and *Eucalyptus*), French Tarragon, spinach, the aquatic plant *Myriophyllum*, and microbial mats. There appeared to be no formal protocol, but the study was described adequately. These results should be treated as qualitative and indicative of trends, not as absolute values, because perchlorate concentrations in the medium were not constant. Rather, they declined with time as the initial addition of perchlorate

was not renewed during the course of the studies, and evidence was provided it was taken up by the plants as well as biodegraded.

Have Limitations in Studies Been Appropriately Discussed? There appeared to be two salient limitations in these tests, and they concern the chronic tests of *Ceriodaphnia* and *Pimephales* and Nzegung's (1998) phytoremediation study. I do not consider the 7-day chronics as definitive because their exposure durations were short, and for one species (*Pimephales*), they did not encompass the species' life cycle. With a chemical that affects hormonal function, I would have liked to have seen a multi-generation test with *Ceriodaphnia* and a full life cycle test with *Pimephales*. The question: were the 7-day test durations sufficient to check on hormonal effects with fish and invertebrates that had no prior exposure to the test substance? There is a lot of uncharted ground with endocrine disrupters, and we cannot yet answer this question. An analysis by Barnthouse, Suter and Bartell (1988) showed that full chronics, i.e., those encompassing the organism's life cycle, were more sensitive than early life stage tests. However, I would not extend the chronic testing unless documented surface water or sediment exposures exceed the 600 µg/L chronic threshold that the Toxicological Review calculated (i.e., Chapter 7, Screening Ecological Risk Assessment for Perchlorate).

The second caveat concerned Nzegung's phytoremediation study. The authors of the Toxicological Review noted the lack of toxicant renewal, and suggested that bioaccumulation probably was underestimated. The lack of perchlorate renewal seems to be a significant design flaw, and it may have underestimated significantly both degradation and bioaccumulation. Future tests should use a renewal or continuous flow technique to better simulate groundwater or irrigation water. The basis for biodegradation should be identified, as Nzegung proposed.

What are the Strengths and Weaknesses of Data Analyses? The data generally were analyzed with appropriate methodologies and to the extent necessary. One important toxicological index, the acute-chronic ratio, was not measured in one set of tests, because of their design. If the BEA tests had measured the 48-hr LC50 for *Ceriodaphnia* and the 96-hr LC50 for *Pimephales*, it would have been possible to calculate acute-chronic ratios for both species. These ratios are essential for estimating chronic toxicity from acute tests, and there are not enough of them available for perchlorate. Although they were calculated by EA for the same species, duplicate values would have afforded a check on precision.

Has the Document Adequately Evaluated the Results of all Relevant Studies and the Biological Significance of the Entire Database? In general, the existing toxicological data were comprehensively presented and interpreted by Sprenger et al. (1998). Including the test results for *Hydra*, *Bufo*, lamprey and Bringman and Kuhn's (1977) *Daphnia magna* test would strengthen the weight of evidence, unless these tests are seriously flawed. The Toxicological Review mentions these tests, but does not include them in the analyses or inferences. Reasons for their exclusion were not given. I believe these tests of other species, some very distant phylogenetically from fish and cladocerans, increase the weight of evidence suggesting that perchlorate probably is not chronically toxic to aquatic life below 600 ug/L.

Completeness of Technical Documentation? The technical documentation was excellent.

Are There Any Sections That Could Be Improved? Because of my concern about ephemeral pond-wetland exposure in arid regions, I think that more testing is needed concerning the following.

Highest Priority	Documentation of perchlorate concentrations in wetlands and small streams in arid regions near areas where significant quantities of perchlorate have been disposed. Small streams and ponds, perennial and ephemeral, in arid regions are very attractive to wildlife because these habitats and water are so limited. Such areas are subject to evaporative concentration. Thus, a combination of limited rainfall, evapo-concentration and bioaccumulation in wetland plants and invertebrates could create potentially higher exposure than assumed thus far.
Highest Priority	Perchlorate's chronic toxicity potential to nesting birds dependent on wetlands (e.g., waterfowl, shorebirds and blackbirds). Chronic tests with appropriate surrogate avian species may be necessary. Rationale: Various birds are very water-dependent for nesting; waterfowl, shorebirds, and blackbirds are often found using ponds and riparian habitat throughout the West. In some locations (San Joaquin, CA; Great Salt Lake, UT), they have sometimes been placed at risk from selenium that has bioaccumulated in their invertebrate food. Wildlife use of ephemeral ponds and perennial ponds subject to extensive evapoconcentration seems to increase their risk. Wildlife using evaporation ponds and wetlands draining areas with significant perchlorate in soils could be at risk, depending on water concentration, magnitude of bioaccumulation, and toxicity.
Highest Priority	Bioaccumulation potential in wetland plant species. Rationale: Bioaccumulation factors of up to 25-fold background were reported in the Toxicological Review, and Nzengung (1998) reported ClO_4 residues in leaves to be 3536 mg/kg in cottonwood, 813 mg/kg in willow, and 641 mg/kg in eucalyptus. Residues in leaves were the highest residues found in plants. These studies should examine bioaccumulation in leaves of woody riparian species, emergent macrophytes, submergent macrophytes, and algae. Species that are known to be important for food of wildlife should be emphasized.
Highest Priority	Perchlorate's chronic toxicity potential to a rodent. The mammalian tests of rodents conducted to support the human health risk assessment will

satisfy this data requirement. Rodents will also be attracted to the wetlands described above.

Highest Priority — Degradation potential in wetland plant species (include microbial role). This study can be dovetailed to the plant bioaccumulation study. It is designed to index persistence.

Medium Priority — Chronic toxicity to fish and amphibia, encompassing species and developmental endpoints that are sensitive to thyroid hormone effects. According to one of the peer reviewers, Dr. Rochelle Tyl, 14-20 days of perchlorate exposure appears to be sufficient to effect hormonal changes in mammals. Therefore, all developmental endpoints should encompass at least a 14-20 day pre-exposure. A long-term chronic exposure with fathead minnows encompassing gametogenesis and embryogenesis up through 28-day old fry is one such study. A study of the effects on tadpole development up through metamorphosis would be helpful in deciding risks to amphibians, as the frog embryo development study (FETAX) performed with perchlorate is inconclusive with respect to hormone-mediated effects.

Low Priority — Sediment toxicity to the freshwater amphipod, *Hyallela*. I accept the Toxicological Review's argument that water column toxicity should be indicative of sediment toxicity, so I believe sediment testing constitutes a low priority. Because of the concern about exposure in evaporative ponds and small wetlands, sediment exposure of invertebrates will occur.

No other recommendations appear to have sufficient priority to mention. I do not believe the evidence suggests risks are sufficient to justify the following studies that were recommended by the Toxicological Review:

Effects on aquatic plants: I saw no indication of phytotoxicity in the studies reviewed. Aquatic plants generally are protected by concentrations that protect aquatic life, because they are less sensitive (Kenaga and Moolenaar (1979).

Effects on nondaphnid invertebrates: Daphnids generally are the most sensitive group of aquatic invertebrates, and given perchlorate's low toxicity, the value of further testing appears questionable.

Effects on litter-feeding invertebrates: Rather than a toxicity study *per se*, examining bioaccumulation in species like corixids (backswimmers) and chironomids (midge flies) – abundant pond-dwelling species - may be more helpful. Some of the corixids and chironomids are tolerant species, and interest in them centers on what perchlorate residues they contain that would provide exposure to birds that eat them.

Testing in estuarine waters: The evidence does not suggest risks to large waterbodies. Groundwater concentrations are too low; groundwater flows are often relatively small relative to overall surface water flows, and dilution capacities are frequently large in large streams and estuaries.

Is the Document Useful for Characterizing Ecotoxicological Effects? If Not, Specify the Nature and Extent of Changes. Yes, the data presented are sufficient to estimate risks, and the Toxicological Review did a good job of estimating risks, from a screening perspective.

SCREENING ECOTOXICOLOGICAL RISK ASSESSMENT (Chapter 7 of Toxicological Review)

Have the Goals and Objectives of the Ecological Screening Analysis Been Adequately Described? Yes. This document was written by scientists well versed in ecorisk assessment and in the interpretation/use of ecotox effect and exposure data.

Does the Analysis Support the Summary and Conclusions Presented? Most of the conclusions are highly supportable.

I am reluctant to fully support the following conclusions/recommendations:

Risks to earthworms (assumed to be representative of soil invertebrates): I do not believe a interspecies safety factor of 242 is warranted, and believe it should be much lower. Strengthening the scientific basis for any safety factor would be helpful. This safety factor resulted in earthworms being categorized as one of the most sensitive species tested, which is inconsistent with the scientific literature on their relative sensitivity.

Risks to herbivorous mammals: The endpoint selected (effect on iodide uptake by the thyroid) has not, to my knowledge, been linked directly to population-level effects. If it is demonstrated to be directly translatable to a population-level effect, then it is acceptable. However, toxicological endpoints for aquatic life and wildlife must be population-level effects. For aquatic life and wildlife, sublethal physiologic or biochemical changes (biomarkers) are not accepted as appropriate surrogates by EPA (Stephan et al. 1985) or the scientific community (Gentile and Slimak1992). Effects on growth, survival and reproductive success are the commonly accepted population-level measurement endpoints.

Are Relevant and Important Aspects of Uncertainty Addressed Sufficiently? Yes. The uncertainty analysis was very thorough, though quite conservative.

Utility of Bioassays for Characterizing Hazard (Effects Potential): The bioassays were acceptable for characterizing hazard, subject to the caveats mentioned above concerning the potential for chronic toxicity over the life cycle of aquatic vertebrates (fish and amphibia). I was

disappointed that the screening ERA only used the *Ceriodaphnia* and fathead minnow acute and chronic data. It is my understanding this action was taken because of concern that the toxicological data for ammonium perchlorate may have reflected ammonia toxicity. Additional analysis might be desirable to see whether any of these data can be used, if it can be shown that ammonia toxicity was not a factor. Including more taxa in the Tier 2 chronic threshold analysis may make it less conservative and representative of more taxa.

Further Ecotoxicological Testing Needs For Perchlorate

Will the Additional Ecotoxicological Studies Currently Underway Be Sufficient to Characterize the Ecotoxicological Potential of Perchlorate? If not, what are data needs and why, and associated experimental designs? The following repeats recommendations made above: Because of my concern about ephemeral pond-wetland exposure in arid regions, I think that more testing is needed concerning the following. Details are provided in the review of Sprenger et al. (1998).

Highest Priority — Documentation of perchlorate concentrations in wetlands and small streams in arid regions near areas where significant quantities of perchlorate have been disposed. Small streams and ponds, perennial and ephemeral, in arid regions are very attractive to wildlife because these habitats and water are so limited. Such areas are subject to evaporative concentration. Thus, a combination of limited rainfall, evapo-concentration and bioaccumulation in wetland plants and invertebrates could create potentially higher exposure than assumed thus far.

Highest Priority — Perchlorate's chronic toxicity potential to nesting birds dependent on wetlands (e.g., waterfowl, shorebirds and blackbirds). Chronic tests with appropriate surrogate avian species may be necessary. Rationale: Various birds are very water-dependent for nesting, and waterfowl, shorebirds, and blackbirds are often found using ponds and riparian habitat throughout the West. In some locations (San Joaquin, CA; Great Salt Lake, UT), they have sometimes been placed at risk from selenium that has bioaccumulated in their invertebrate food. Wildlife use of ephemeral ponds and perennial ponds subject to extensive evapoconcentration seems to increase their risk. Wildlife using evaporation ponds and wetlands draining areas with significant perchlorate in soils could be at risk, depending on water concentration, magnitude of bioaccumulation, and toxicity.

Highest Priority — Bioaccumulation potential in wetland plant species. Rationale: Bioaccumulation factors of up to 25-fold background were reported in the Toxicological Review, and Nzengung (1998) reported ClO_4 residues in leaves to be 3,536 mg/kg in cottonwood, 813 mg/kg in willow, and 641

mg/kg in eucalyptus. Residues in leaves were the highest residues found in plants. These studies should examine bioaccumulation in leaves of woody riparian species, emergent macrophytes, submergent macrophytes, and algae. Species that are known to be important for food of wildlife should be emphasized.

Highest Priority — Perchlorate's chronic toxicity potential to a rodent. The mammalian tests of rodents conducted to support the human health risk assessment will satisfy this data requirement. Rodents will also be attracted to the wetlands described above.

Highest Priority — Degradation potential in wetland plant species (include microbial role). This study can be dovetailed to the plant bioaccumulation study. It is designed to index persistence.

Medium Priority — Chronic toxicity to fish and amphibia, encompassing species and developmental endpoints that are sensitive to thyroid hormone effects. According to one of the peer reviewers, Dr. Rochelle Tyl, 14-20 days of perchlorate exposure appears to be sufficient to effect hormonal changes in mammals. Therefore, all developmental endpoints should encompass at least a 14-20 day pre-exposure. A long-term chronic exposure with fathead minnows encompassing gametogenesis and embryogenesis up through 28-day old fry is one such study. A study of the effects on tadpole development up through metamorphosis would be helpful in deciding risks to amphibians, as the frog embryo development study (FETAX) performed with perchlorate is inconclusive with respect to hormone-mediated effects.

Low Priority — Sediment toxicity to the freshwater amphipod, *Hyallela*. I accept the Toxicological Review's argument that water column toxicity should be indicative of sediment toxicity, so I believe sediment testing constitutes a low priority. Because of the concern about exposure in evaporative ponds and small wetlands, sediment exposure of invertebrates will occur.

LITERATURE CITED

Barnthouse, L.W., G.W. Suter II, and S.M. Bartell. 1988. Quantifying risks of toxic chemicals to aquatic populations and ecosystems. Chemosphere 17(8):1487-1492.

Gentile, J.H. and M.W. Slimak. 1992. Endpoints and indicators in ecological risk assessments. pp. 1385-1397. In: [D.H. McKenzie, D.W. Hyatt, and V.J. McDonald, Editors]. Ecological Indicators. Elsevier Applied Science, New York.

Kenaga, E.E. and J. Moolenaar. 1979. Fish and Daphnia toxicity as surrogates for aquatic vascular plants and algae. Environmental Science and Technology. 13:1479-1480.

Stephan, C.E., D.I. Mount, D.J. Hansen, J.H. Gentile, G.A. Chapman, and W.A. Brungs. 1985. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. U.S. Environmental Protection Agency, Washington, D.C. NTIS No. PB85-227049. 98 pages.

3.9 Risk Analysis: Dr. Melvin Andersen

Summary

This document develops a rationale for a mode of action based approach to assessing noncancer and cancer risks of perchlorate and also pursues a conventional assessment calculating an RfD from a specific critical study with application of a series of uncertainty factors. Unfortunately, the mode of action data do not appear to have much quantitative influence on the RfD determinations. Many of the studies upon which this hazard characterization is based are in the process of completion and several important recommendations for further analysis of the results were provided at the review. These new data should be forthcoming over the next several months and will almost assuredly influence the final RfD. It appears premature to offer a firm number in this current hazard characterization document. The chances are great that any provisional RfD will have to be changed within the next 6 to 9 months. While such a provisional standard could be set today, it would be better to postpone promulgation of a single number until the data set is completed and analyzed. An alternative would be to offer a range of RfD values depending on choices of LOAELs, LOELs, BMDs, and uncertainty factors.

A clearer articulation of the proposed strategy for using mode of action, pharmacokinetic, and iodide uptake inhibition data quantitatively should be added to the document either in the text or as an Appendix. This material would allow a reader to see clearly how the mode of action data are expected to influence the quantitative hazard characterization with perchlorate. The harmonization of cancer and non-cancer assessments proposed is appropriate, although not all the issues in the use of the data for conducting a cancer margin of exposure (MOE) are well-developed. Finally, some reorganization of the document would be helpful for clarity. Despite some limitations, this document is important in outlining the current status of the science for a mode of action assessment with perchlorate. It collates a large body of information, points the direction for a comprehensive assessment based on mode of action, and provides interim guidance while the data are collected, analyzed, and published. The US EPA staff and their collaborators are to be commended on the quality of this interim guidance for a perchlorate hazard characterization.

General

In contrast to other risk characterization documents from US EPA, this document is unusual. It actually represents a work in progress more than a completed characterization of health and ecological risks posed by perchlorate. Several studies are only partially reported (for instance, the immunotoxicity evaluations), and the development of a more integrated evaluation of the multiple toxic endpoints based on mode-of-action of perchlorate as an inhibitor of iodine uptake by the thyroid is outlined, but only partially implemented. The text in some places is uneven, for instance the extensive evaluation/BMD calculations conducted on the various data sets, given the final position that the thyroid changes in the post-natal day 5 (PND5) pups serves as the single response for estimating a provisional RfD. In addition, the placement in the text of the correlation testing with T4, T3, TSH and thyroid histology detracts from the flow of the document. These evaluations could be included as appendices with the main conclusions of the analyses captured in the body of the hazard characterization document.

The US EPA staff and collaborators have not simply provided a document based on conclusions drawn from other primary work and created a secondary document with interpretation. Instead, they have proactively participated in design of a research/testing program and the analysis and interpretation of the results. The statistics and correlations are a necessary part of the whole. With the new cancer guidelines, it will be important that the EPA conduct several mode of action based risk assessments in order to take an informed position on the manner in which these new guidelines should be implemented. This example with perchlorate is the first such effort by the US EPA staff. While there is room for improvement and clarification in the document, as noted below, US EPA staff deserves to be commended for taking this important first step under the new guidelines.

The mode of action based approach with perchlorate is important for assessing significance of hormonal changes and thyroid hypertrophy as a sentinel precursor for all relevant responses to perchlorate. The arguments for this approach are included in various portions of the text, although not clearly emphasized in the summary. This approach represents a move in the direction of harmonization of non-cancer and cancer risk assessment paradigms and permits better (i.e., more rationale) use of available biological and toxicological data. The presentation of the benchmark dose calculations points out the difficulties faced in fitting various models to the observed response data. After this fitting is completed, which model form should be applied and how confident are you that you have selected an appropriate mode to represent the data? The section on curve fitting and BMD calculations was the most difficult one to follow. Unfortunately, a more comprehensive biological model of impaired thyroid function for assessing curve shapes and assessing the significance of alterations in hormone concentrations has not yet been completely fleshed out, although the data for such an enterprise are apparently available or being collected. The presentation by Ms. Annie Jarabek (US EPA) included some detail about these pharmacokinetic and mechanistic studies. However, the manner in which the perchlorate pharmacokinetics and iodide inhibition data will be used in the final hazard characterization was never clearly articulated. The document could be improved by adding a

section that explained the quantitative impact of knowledge mode of action in the risk assessment.

Specific Comments

Most of the studies were adequately described and the interpretations were easily followed. However, the presentation and rational for the immunotoxicology studies were not well presented and the labeling of these studies by letter designation in the report was uninformative. It was also unclear whether any of these planned immunology studies will help at all in considering the importance of the human experience where altered hematology was observed in patients treated with perchlorate. The document states on page 3-6 that these hematological effects were believed to be due to an immunological response.

One of the strengths of the document was in conveying the results of the various thyroid hormone studies in tabular form. This presentation included the relationships for the thyroid hormone, TSH, and histology. These tables present a large amount of material in a readily comprehensible fashion. A tabulation of effect levels for specific toxicity in tissues other than the thyroid would also be useful. The presentation of the associations between T3, T4, TSH, and pathological changes was very important for confirming a link between the several steps: inhibition of I_2 uptake, impaired synthesis/release of T3 and T4, feedback increase of TSH, and overt histological alterations. This section should be improved by including an easily followed preamble letting the reader know exactly what is being done and why it is being done.

There are a number of sections that could be improved by some changes in organization and emphasis. In general, the text would be clearer if the specifics of the statistical analyses were provided in an appendix, with the results and implications emphasized in the text itself. The document has a conventional RfD assessment portion, i.e., critical study and application of uncertainty factors. It also contains a nascent structure in which mode of action is used to categorize effects and provide a larger structure for interpretation of studies. Some data are interpreted in terms of adequacy of a particular study to assess a NOAEL and others in terms of consistency, with the hypothesis that toxicity is secondary to precursor thyroid effects. The manner in which these different points are made and emphasized probably deserves more attention by the authors. This comment is not so much a criticism as point of emphasis. This presentation mode of action is relatively new for a non-cancer assessment. The authors need to think about how this emphasis might lead to reorganization of the presentation of materials. A possibility is to add text to the Executive Summary that lets the reader know what is coming in terms of the mode of action emphasis.

The document adequately captures the current state of knowledge for perchlorate toxicity. The new toxicity tests cover the concerns raised in earlier evaluations of the adequacy of the data base. One drawback is that many of the newer studies have not yet been adequately prepared, reviewed, and documented. All these results should be published and undergo peer-review separate from this evaluation. The argument supporting a precursor relationship for thyroid

dysfunction is well organized. However, this hypothesis is not implemented in a way that can have much quantitative impact on the assessment. If a quantitative implementation of the mode of action is planned, it would be useful to convene a panel either constituted internally at US EPA or jointly with external consultants to ask questions about what form the quantitative approach will take and whether the appropriate kinetic and pharmacodynamic data will be available to provide confidence in a more quantitative approach to guide the risk assessment.

Because of the mode of action emphasis, there is a tension between normal use of a single study and some consideration of the entire body of data in the determination of the RfD. There are two points of concern here. One is how the mode of action should influence the consideration of the entire data set; the second is the estimation of the toxic endpoints most sensitive to disruption of thyroid hormone homeostasis. It probably is appropriate and necessary to establish the most sensitive endpoint; however, depending on mode of action, the most sensitive endpoint in a rodent model may not turn out to be the most sensitive endpoint expected in humans (Barton and Andersen, 1998). The mode of action and evaluation of the totality of the data base might be more directly helpful in the selection of appropriate uncertainty factors after a critical study or several critical studies are identified. In addition, a quantitative model of disruption of thyroid hormone homeostasis (i.e., a BBDR model) might actually render some of the uncertainty factors unnecessary.

As noted earlier, it maybe necessary to select several critical studies. The critical study will depend on presumed mode of action and on characterization of animal-human differences in sensitivity of all organ systems. These targets include the target thyroid itself and the systemic tissues that respond to the thyroid hormone signals. The use of the PND5 results, with thyroid tissues as the critical study, seems sound given the present knowledge. However, it may not be chosen after all the new data are collected and ancillary studies of these tissues are completed, as suggested at the review. One concern is the consequences of the observed changes in the pup thyroid. Do they lead to adverse outcomes in these rats as they mature or are they simply a compensatory effect related to a temporary impairment in iodide incorporation that will be quickly rectified during nursing and weaning? These questions may require other assessments in the adult rats for altered behavior, etc. Also, are the same effects expected in humans? And, are there situations that make humans more sensitive or less sensitive at some point in life than the rat? These questions are all relevant to the selection and use of this observation as a minimal LOAEL. The Panel was not in agreement about the 'adversity' of the changes in thyroid in these rats and whether they represent a minimal LOAEL or simply a LOEL.

An uncertainty factor of 3.0 was used for animal-human differences because of pharmacodynamic differences in sensitivity across species. Given (1) the care taken to show the correlation of the various endpoints for thyroid effects, (2) the lack of finding suggesting other modes of action for perchlorate, and (3) the differences in human and rat plasma reserves of thyroid hormones, a factor of 3.0 seems appropriate at this time for this particular endpoint. It could be further adjusted by use of a PK/PD model for perchlorate kinetics/iodide uptake inhibition/ thyroid hormone alterations resulting from perchlorate. Further work on the kinetics

of perchlorate in different species and the kinetics of inhibition of the iodide transport mechanisms by perchlorate in rat and human thyroid tissues will be useful in confirming the assumptions above. One point that deserves emphasis throughout the text is the importance of mode of action in assessing animal-human differences with induced thyroid hormone suppression. It is clear that rats are more sensitive to promoting effects on the thyroid leading to carcinogenesis. However, the obverse side of this argument is the question of the relative sensitivity of rats and humans to toxic effects related to a hypothyroid condition arising from perchlorate exposure. This distinction of known species differences for cancer and the lack of knowledge of species-differences for non-cancer effects needs to be clearly articulated in this section dealing with species differences.

Three other uncertainty factors were applied, each with a value of 3.0, for using a minimal LOAEL versus LOAEL/intrahuman variability, data base deficiencies, and intrahuman variation (sensitive populations). All the corrections together provide a composite correction of 300. This is probably sufficient (and perhaps overly conservative) given the data because of the knowledge of the mode of action, the belief that rats are more sensitive than humans, and the evaluation of minimal effects in the postnatal rats. The number is large enough in composite; however, it is difficult to assert that the factors that constitute the 300 are correctly parsed between the individual factors. The data base and minimal LOAEL uncertainty factors are accessible to experiment, as is the term for interspecies extrapolation above. The Panel was reluctant to call the PND5 changes adverse and suggested further evaluations of these tissues before deciding whether the result should be called a minimal LOAEL or simply a NOEL. In general, I become concerned in reducing the sensitive subpopulation factor from 10, unless specific factors can be cited in support of the change for hypothyroid states. Nonetheless, the total value of 300 does appear health-protective at this time.

The development of a harmonized cancer/non-cancer assessment is a welcome direction within EPA. While not yet completed in the document, the steps taken in this direction are clear and appropriate for a harmonized approach. However, cancer and non-cancer assessments have some differences that still have not been reconciled. Unless they are discussed and reconciled, the authors should probably use different studies and different low dose extrapolation strategies for cancer and non-cancer effects. The present differences can be briefly outlined. First, mode of action has been more frequently (although still very infrequently) applied to cancer risk assessments. In the new US EPA carcinogen guidelines, there is provision to discuss mode of action and to evaluate the possibility that common modes of action could underlie both cancer and non-cancer endpoints. Nasal toxicity with vinyl acetate is an example. In the present case, the cancer assessment for compounds disrupting thyroid hormone homeostasis has to be extended to consider a much broader set of endpoints than simply carcinogenesis.

With the emphasis on MOE approaches for carcinogens with non-linear modes of carcinogenic action, it is likely that certain non-cancer endpoints will become the limiting or critical effects, not the cancer. This situation is likely to be true with perchlorate. For cancer, however, the use of the PND5 studies to assess lifetime cancer risk does not appear optimal.

Alterations in thyroid hormone function throughout life would probably be a better precursor step for a cancer risk assessment. (Emphasis in cancer pathogenesis on measures of disruption over a longer time argues for a cancer risk assessment based on thyroid hormone alterations occurring in adults rather than those occurring at a single time point in neonatal rats.) The authors should consider calculating a BMD for thyroid disruption in adult rats and follow a non-linear cancer risk assessment paradigm to arrive at a cancer-based RfD. This alternative exercise would provide a check on the consistency of conclusions reached for cancer and for the critical non-cancer effects. A problem with harmonizing cancer and non-cancer risk assessments at the present time is the US EPA requirement that non-cancer assessments decide on values of uncertainty factors, while the non-linear cancer assessment estimates a margin of exposure (MOE) by comparing a dose-adjusted BMD with presumed human exposure levels. The non-linear cancer assessments beg the question of how large does the calculated MOE have to be for the risk assessor to be comfortable. Somehow any harmonized assessment has to reconcile these two disparate endgames – uncertainty factors versus MOEs. Despite this unresolved issue of different quantitative strategies having to be applied to cancer and non-cancer assessments, this document has come closer to achieving a harmonized human health risk assessment than has been done for any other compound in the United States.

The utility of further studies depends partially on decisions regarding adequacy of the studies supported over the past few years. With the exception of the immunotoxicology, the test appeared to follow standard protocols accepted by the toxicology community and by EPA. The immunotoxicology requirements and the interpretation of these tests are both still evolving. Some standard immunotoxicity tests such as the sheep red blood cell antibody tests were not performed, and the studies did not appear to be tailored to create special test protocols based on thyroid alterations. In general, it was surprising that the design criteria used for most of the new studies didn't include a review of expected non-cancer endpoints for other iodide uptake inhibitors and considerations of modifying protocols based on that knowledge. (This may well have been a consideration; it's just not apparent from the materials reviewed.) Some other tests might still be considered, including evaluation of other developmental delays or of permanent impairment in adult animals following fetal/neonatal exposures. Have potential endpoints been adequately examined to rule out higher sensitivity of the fetus to transient periods of hypothyroidism?

There is a potential for considerable 'value added' in the assessment by development of a BBDR model for iodide uptake inhibition, thyroid hormone depletion, hypothyroidism, and compensation. Qualitatively, defining mode of action would assist in establishing non-linear approaches for cancer risk assessment and in supporting choices for uncertainty values. However, the analysis of dose response curves with perchlorate relied on external exposure concentrations and on curve fitting to forms that are not concordant with specific biological processes. The development of a PBPK model should permit a more quantitative approach to deriving uncertainty factors, to 'fitting' biological models with clearly defined parameters to data, and to inferring the impact of differences between humans and rodents for the risk assessments. What should this 'model' look like? It is important that the design of the

biologically based risk model begin as soon as a hypothesis is elaborated. In other words, it should be well along the development trail already.

One of the most important criterion in the modeling is a decision about which endpoint should be evaluated. This decision leads to the objective function against which model performance is to be judged. Thus it becomes important to define the measure of response or adversity that will be the basis for regulation - is it some proportionate change in circulating T4/T3, some alteration in TSH, or some characteristic alteration in the cellular structure of the thyroid tissue? The material presented in the document does not provide an adequate basis for commenting on any current model structure, although the methods for the PK, iodide uptake, and thyroid homeostasis models are fairly conventional and efforts in linking these individual elements have been previously reported.

There is presently inadequate documentation of the data that would be collected for the development of a quantitative model. However, the development of such a quantitative model represents the primary tool needed to move from qualitative application of the broader array of mode of action data to its quantitative impact on risk assessments. The approach needed is the organization of a biologically based dose response model as described in the cancer risk assessment guidelines. These models influence analysis in the range of observation and enlighten the extrapolation between species and from high to low doses. It may be too late in the effort to complete the perchlorate review to have a BBDR model that influences the risk assessment in any quantitative fashion. Nonetheless, pursuing such a model to completion now could give more confidence in the correctness of any decisions made in the short-term regarding perchlorate and, more importantly, guide future efforts with harmonized mode of action based assessments.

Some specific comments

The analogies to parking brakes and hillside parking were uninformative and confusing (see, page 2-8). Simply stated perchlorate is stable in biological matrices. That wording should be sufficient.

page 3-13 One of the potassium chlorates, third line up from the bottom, has the potassium as the labeled isotope vice the chlorine.

page 3-14 Perchlorate reduced the thyroid blood ratio of what? Iodide, I presume, although the sentence could be read to indicate perchlorate.

page 3-15 Is it true that the concentration of perchlorate in the thyroid was inversely related to dose, or is the concentration expressed as percent of administered dose related inversely to dose. The former seems hard to reconcile to the behavior of dose-dependent systems.

page 5-18 biological significance is .....suspect. Do you mean .....is uncertain?

page 5-41 The discussion of U-shaped dose response curves probably deserves mention later in the document when discussing limiting toxicity for different endpoints. Is it real? If it were used as an effect level, does it change the perception of which study gives the lowest adverse effect level? Should anything be done to pursue this observation?

page 5-54 Immunotoxicity – is there any data on immune function in hypothyroid individuals from which hypotheses could be generated about the endpoints that might be most worthwhile to pursue.

page 6-41 exquisite or equisite.

REFERENCES

Barton, H.A., Andersen, M.E., and Clewell, H.J., III (1998). Harmonization: Developing Consistent Guidelines for Applying Model of Action and Dosimetry Information to Cancer and Noncancer Risk Assessment. *Hum. Ecol. Risk Assess 4: 75-115.*